

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw. model

Run on: April 29, 2001, 06:40:29 ; Search time 14.58 seconds

(without alignments)  
608.042 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678

Sequence: 1 MSPPKYPSSLRTPETLPD.....DRREKLQEGKLDRTFHLST 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.67:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	2 JE0383	NADH dehydrogenase
2	514	75.8	129	2 S28237	NADH dehydrogenase
3	315	46.5	133	2 JC2003	NADH ubiquinone ox
4	149	22.0	179	2 A33164	hypothetical prote
	88.5	13.1	1078	2 S77162	DNA topoisomerase
7	74	10.9	308	2 G83104	probable ferredoxi
	74	10.9	451	2 F75177	tryptophan synthas
8	70.5	10.4	788	2 S75209	dnax protein - Syn
9	70	10.3	695	2 G64327	H+-transporting AT
10	70	10.3	1025	2 T18376	multidrug resistan
11	69.5	10.3	294	2 S22440	protein kinase (EC
12	69	10.2	298	2 S69523	hypothetical prote
13	69	10.2	1023	2 S12519	glutactin - fruit
14	69	10.2	1660	2 T18561	vitellinogen vit-6
15	68.5	10.1	294	2 S23095	protein kinase (EC
16	68.5	10.1	574	2 T49271	CELL DIVISION CONT
17	68.5	10.1	294	2 S57072	hypothetical prote
18	68.5	10.1	1330	2 A36373	hypothetical prote
19	68.5	10.1	1333	2 T38401	retrotransposabl
20	68	10.0	185	2 S12205	hypothetical prote
21	68	10.0	716	1 JC5061	macrophage-stimula
22	67.5	10.0	133	1 MNWQBY	genome-linked prot
23	67.5	10.0	289	2 E75391	conserved hypothet
24	67.5	10.0	296	2 C82645	transcription regu
25	67.5	10.0	600	2 C83142	hypothetical prote
26	67	9.9	194	2 A82072	conserved hypothet
27	67	9.9	232	2 B75121	hypothetical prote
28	67	9.9	316	2 T21180	hypothetical prote
29	67	9.9	455	2 S71344	putnergic recepto

30	66.5	9.8	472	2 B37777	methyl viologen-re
31	66.5	9.8	674	2 E83263	hypothetical prote
32	66.5	9.8	634	2 S32230	Ca2+-transporting
33	66.5	9.8	1036	2 A57386	preprotein translo
34	66	9.7	193	2 H81874	probable outer mem
35	66	9.7	271	2 T24965	hypothetical prote
36	66	9.7	342	2 D71324	conserved hypothet
37	66	9.7	353	2 F83570	hypothetical prote
38	66	9.7	413	2 H75357	tRNA (5-methylamin
39	66	9.7	657	1 FOVZEV	major core protein
40	66	9.7	677	2 S54561	RNA14 protein - ye
41	65.5	9.7	354	2 T22108	hypothetical prote
42	65.5	9.7	899	2 B38529	nib protein - Esc
43	65.5	9.7	959	2 D83339	glycine cleavage s
44	65.5	9.7	2358	2 T39569	probable alpha-glu
45	65.5	9.7	2371	2 T43432	alpha-glucan synth

## ALIGNMENTS

RESULT 1  
JE0383  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain NDUF84 - human  
N:Alternate names: NADH:ubiquinone oxidoreductase subunit NDUF84  
C:Species: Homo sapiens (man)  
C>Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 21-Jul-2000  
R:Loeffen, J.L.C.M.; Triepels, R.H.; van den Heuvel, L.P.; Schuelke, M.; Buskens, C.A.  
Biochem. Biophys. Res. Commun. 253, 415-422, 1998  
A>Title: cDNA of eight nuclear encoded subunits of NADH:ubiquinone oxidoreductase: Hu  
A:Reference number: JE0379; MUID:99097250  
A:Accession: JE0383  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-129 <LOE>  
A:Cross-References: EMBL:X64898; NID:94164445; PIDN:AAD05421.1; PID:94164446  
C:Keywords: NAD; oxidoreductase

Query Match 100.0%; Score 678; DB 2; Length 129;  
Best Local Similarity 100.0%; Pred. No. 6; le-65;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MSPPKYPSSLRTPETLPDAEYNISPEPTRRAOERLARAQKREYLLQYNDPNRGLI 60  
DB 1 MSPPKYPSSLRTPETLPDAEYNISPEPTRRAOERLARAQKREYLLQYNDPNRGLI 60  
OY 61 ENPALRMAYARTINYPNFRPTPKNSLMGALCGFGLIFYYIIKTERDRREKLQEGK 120  
DB 61 ENPALRMAYARTINYPNFRPTPKNSLMGALCGFGLIFYYIIKTERDRREKLQEGK 120  
OY 121 LDRTFHLST 129  
DB 121 LDRTFHLST 129  
RESULT 2  
S28237  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain B15 - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 05-Nov-1999  
A:Accession: S28237  
R:Walker, J.E.; Arizumi, J.M.; Dupuis, A.; Fearnsley, I.M.; Finel, M.; Medd, S.M.; P  
J. Mol. Biol. 226, 1051-1072, 1992  
A>Title: Sequences of 20 subunits of NADH: ubiquinone oxidoreductase from bovine hear  
A:Reference number: S28237; MUID:92389317  
A:Accession: S28237  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-129 <WAL>  
A:Cross-References: EMBL:X64898; NID:94113; PIDN:CAA46107.1; PID:94114  
C:Keywords: electron transfer; mitochondrion; NAD; oxidoreductase





Query Match 10.3%; Score 70; DB 2; Length 1025;  
Best Local Similarity 33.3%; Pred. No. 35;  
Matches 14; Conservative 11; Mismatches 17; Indels 0; Gaps 0;

OY 85 KNSLMGALCGFGLFIYIIRTERDRKELIQEKLDRFTF 126  
Db 500 KNSLGSVLEIGLITLYSTIKITKWKRRKIRKANEMDNVH 541

RESULT 11

S22440  
Protein kinase (EC 2.7.1.37) cdc2 homolog 1 - rice

C:Species: Oryza sativa (rice)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 18-Jun-1999

C:Accession: S22440

R:Hashimoto, J.; Hirabayashi, T.; Hayano, Y.; Hata, S.; Ohashi, Y.; Suzuki, I.; Utsugi, M.; Gen. Genet. 233, 10-16, 1992

A:Title: Isolation and characterization of cDNA clones encoding cdc2 homologues from Ory

A:Reference number: S22440; MUID:92293101

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-294 <HAS>

A:Cross-references: EMBL:X60374; NID:g20342; PIDN:CAA42922.1; PID:g20343

C:Superfamily: kinase-related transforming protein; protein kinase homology

C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein

F:2-56/Domain: protein kinase homology <IN>

F:10-18/Region: protein kinase ATP-binding motif

F:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 10.3%; Score 69.5; DB 2; Length 294;  
Best Local Similarity 36.2%; Pred. No. 8.3;  
Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

RESULT 12

S69523  
Hypothetical protein 17 - phage HPI

C:Species: phage HPI

C:Date: 06-Dec-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Oct-1999

C:Accession: S69523

R:Fitelson, D.; Fitzmaurice, W.P.; Benjamen, R.C.; Goodman, S.D.; Waldman, A.S.; Scocca, N.

A:Title: The complete nucleotide sequence of bacteriophage HPI DNA.

A:Reference number: S69503; MUID:96279738

A:Accession: S69523

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-298 <ESP>

A:Cross-references: EMBL:U24159; NID:g1046235; PIDN:AA09202.1; PID:g1046244

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995

Query Match 10.2%; Score 69; DB 2; Length 298;  
Best Local Similarity 26.9%; Pred. No. 9.5;  
Matches 39; Conservative 14; Mismatches 50; Indels 42; Gaps 8;

OY 11 LRTPLTDLDAEY--NISPETRR---AOAERLAIRAQLKRE-YLLOYNDPNRGLIENP 63  
Db 30 LHEAETVDEHYANLMPERHFNMGQVTELKAEENKGETOLPAIAPNK-LIEYN 88  
OY 64 ALLMAVARTINVPNFRPPKNSL-----MGALCGFGLPI 99  
Db 83 RAGQYLTFT-SIELTPNFRNSGKAVLSGLVTDSPASVGTTELKFNNAQSGVSG----- 142

OY 100 FIYIITTEDRKELIQEKLDRFT 124

Db 143 ----EFIKVDSAKED-VEEEKALRT 163

RESULT 13

S12519  
Glutactin - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Jul-2000

C:Accession: S12519

R:Olson, P.F.; Fessler, L.I.; Nelson, R.E.; Sterne, R.E.; Campbell, A.G.; Fessler, J.

A:Title: Glutactin, a novel Drosophila basement membrane-related glycoprotein with se

A:Reference number: S12519; MUID:90214632

A:Accession: S12519

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1023 <OLS>

A:Cross-references: EMBL:X53286; NID:g297084; PIDN:CAA37380.1; PID:g297085

C:Genetics: A:Introns: 390/3

Query Match 10.2%; Score 69; DB 2; Length 1023;  
Best Local Similarity 26.5%; Pred. No. 44;  
Matches 36; Conservative 15; Mismatches 45; Indels 40; Gaps 8;

OY 2 SFPKPPSSLTLPETLDPAEYNISPETRAQAQ---RLAIRAQLKREYLLQY-----N 52  
Db 790 SYEYGPENGNLEPET--DANRNFSEEDREQOQEQQLRREQOQEQREYQLQREDOER 847

OY 53 DPNRGLIE-NPALLRMAVARTINVPNFR-----TPKNSLMGALCGFGLFIYIYIK 106  
Db 848 EOERGOQEPGP-----EYPSYEEYSRALQEKNAERDRI-----YAE 886

OY 107 TERDR--KEKLIOECK 120  
Db 887 QERERQOQETLQENQ 902

RESULT 14  
S18561  
Vitellogenin vit-6 [similarity] - Oscheilus sp. (strain CEM1)

C:Species: Oscheilus sp.

A:Variety: S01131

C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 01-Dec-2000

C:Accession: T18561

R:Winter, C.E.; Penha, C.; Blumenthal, T.

A:Title: Comparison of a vitellogenin gene between two distantly related rhabditid ne

A:Reference number: Z18974; MUID:96212969

A:Accession: T18561

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1660 <WIN>

A:Cross-references: EMBL:U05449; NID:g1515336; PID:g1515337; PIDN:AA049749.1

A:Genetics: A:Note: CEM1-vit-6

A:Superfamily: vitellogenin

Query Match 10.2%; Score 69; DB 2; Length 1660;  
Best Local Similarity 24.8%; Pred. No. 81;  
Matches 34; Conservative 20; Mismatches 57; Indels 26; Gaps 4;

OY 6 YKPSLTLPETLDPAEYNISPETRAQAQRLAIRAQLKREYLLQYNDPNRGLIENPAL 65  
Db 631 YTWSTLTKTISESNPAE---KEIRRVSSGLASTIVEEKKYLESKHKTFNFMOSGAT 686  
OY 66 LRMAVARTINVPNFRPPKNSLMGALCGF-----PLIFLY-----YIKTE 108  
Db 687 LNMA-----TIFNSDVLPRFETASLETYVGGEMNKYLAQIGLYONNDSVLSKLQAYE 741

QY	109	RDRKEKLIQEGKLDRTF	125
		: :      :	
Db	742	ETGLEQLVVRGKRSSSF	758

## RESULT 15

protein kinase (EC 2.7.1.37) cdc2 - Arabidopsis thaliana  
N:Alternate names: cdc2 protein homolog; CDC2a protein; cell division control protein 2  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text\_change 20-Jun-2000  
C:Accession: S23095; A48984; J01337; J00967; S18202  
R:Imajuku, Y.; Hirayama, T.; Endoh, H.; Oka, A.  
FEBS Lett. 304, 73-77, 1992  
A:Title: Exon-intron organization of the Arabidopsis thaliana protein kinase genes CDC2a  
A:Reference number: S23095; MUID:92316202  
A:Accession: S23095  
A:Molecule type: DNA  
A:Residues: 1-294 <IM>  
A:Cross-references: EMBL:DI0850; NID:9217848; PIDN:BA01623.1; PID:9217849  
R:Imajuku, D.; Ferreira, P.; Hemerly, A.; Van Montagu, M.  
Biochem. Soc. Trans. 20, 80-84, 1992  
A:Title: Control of cell division in plants.  
A:Reference number: A48984; MUID:92339744  
A:Accession: A48984  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-294 <IM>  
A:Experimental source: flower  
A:Note: sequence extracted from NCBI backbone (NCBI:P:109461)  
R:Hirayama, T.; Imajuku, Y.; Anal, T.; Matsui, M.; Oka, A.  
Gene 105, 159-165, 1991  
A:Title: Identification of two cell-cycle-controlling cdc2 gene homologs in Arabidopsis  
A:Reference number: J01337; MUID:92039027  
A:Accession: J01337  
A:Molecule type: mRNA  
A:Residues: 1-294 <HF>  
A:Cross-references: EMBL:X57839; NID:916218; PIDN:CAA00971.1; PID:916219  
R:Ferreira, P.C.G.; Hemerly, A.S.; Villarroel, R.; Van Montagu, M.; Inze, D.  
Plant Cell 3, 531-540, 1991  
A:Title: The Arabidopsis functional homolog of the p34cdc2 protein kinase.  
A:Reference number: J00967; MUID:93005715  
A:Accession: J00967  
A:Molecule type: mRNA  
A:Residues: 1-294 <FER>  
A:Cross-references: GB:S45387; NID:9257373; PIDN:AAB23643.1; PID:9257374  
A:Experimental source: flower  
A:Comment: The protein is a key component of the eukaryotic cell cycle.  
C:Keywords: cdc2  
A:Title: cdc2  
A:Accession: S23095  
A:Residues: 3/3; 65/2; 105/3; 163/3; 218/2; 247/3; 265/3  
C:Function:  
A:Description: phosphotransferase; protein kinase; required for G1 to S-phase transition  
C:Superfamily: kinase-related transforming protein; protein kinase homology  
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protein kinase  
E:2-256/Domain: protein kinase homology <KIN>  
E:10-18/Region: protein kinase ATP-binding motif  
E:33,51,127,129/Active site: Lys, Glu, Asp, Lys #status predicted

	Query Match	Similarity	Best Local	Score	DB	length
	Matches	17	Conservative	10.1%	68.5	294
				Pred	No. 11	
				Mismatches	21	Gaps
				Indels	1	1
QY	2	SEPPKKPSLRTPETLPDAEYN-ISPETRRQAQERLAIRAQIKREY	47			
DB	240	APPKMPDLEFVPLDDGVGLSKMLMDPTKRIINRAALEHY	286			



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2001, 06:45:14 ; Search time 9.22 Seconds  
(without alignments)  
479.280 Million cell updates/sec

Title: US-09-726-899-3  
Perfect score: 678  
Sequence: 1 MSPPKPKSSLRTPETLDP.....DRKELIOEGKLDRTFLSY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

T number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	673	99.3	128	1	095168 homo sapien
2	509	75.1	128	1	P48305 bos taurus
3	315	46.5	133	1	P48306 gallus galli
4	88.5	13.1	1078	1	P77966 synecocyst
5	71	10.5	270	1	0933x1 zymomonas m
6	70.5	10.4	771	1	P73098 synecocyst
7	70	10.3	384	1	P35713 homo sapien
8	70	10.3	695	1	Q57675 methanococ
9	69.5	10.3	294	1	P29618 oryza sativ
10	69	10.2	298	1	P51719 bacterioph
11	69	10.2	658	1	P42898 homo sapien
12	69	10.2	1023	1	P33438 drosophila
13	69	10.2	1660	1	Q94637 oscheilus dr
14	68.5	10.1	294	1	CC2A_ARATH
15	68.5	10.1	574	1	VJ23_YEAST
16	68.5	10.1	1333	1	RDPO_SCHPO
17	68	10.0	270	1	LPSC_RHIME
18	67.5	10.0	153	1	VRG_BYDVP
19	67	9.9	455	1	P2X5_RAT
20	67	9.9	618	1	M3K2_HUMAN
21	66.5	9.8	1036	1	SECA_SPIOL
22	66	9.7	193	1	LOLB_NEIMA
23	66	9.7	657	1	VP4B_FOMPV
24	66	9.7	677	1	RN14_YEAST
25	65.5	9.7	450	1	DHE4_LACBI
26	65.5	9.7	2358	1	MOKD_SCHPO
27	64.5	9.5	212	1	DSEA_BURCE
28	64.5	9.5	315	1	RSEB_HAEIN
29	64.5	9.5	654	1	P4792 haemophilus
30	64.5	9.5	842	1	MYH7_MOUSE
31	64.5	9.5	1003	1	ATC_YEAST
32	64.5	9.5	1247	1	NOS_ANOST
33	63.5	9.4	294	1	CC2_MAIZE

34	63.5	9.4	461	1	MANA_EMENI
35	63.5	9.4	686	1	KNLC_STRPU
36	63.5	9.4	878	1	SYA_BACSU
37	63	9.3	188	1	YB69_HAEIN
38	63	9.3	285	1	YN26_YEAST
39	63	9.3	400	1	Y137_CAEEL
40	63	9.3	424	1	KASA_STRCO
41	63	9.3	721	1	CIAT_DROME
42	62.5	9.2	587	1	U184_HCMVT
43	62.5	9.2	966	1	AMPN_FELCA
44	62.5	9.2	1521	1	EM5_CAEEL
45	62.5	9.2	1603	1	V175_CAEEL

## ALIGNMENTS

RESULT	ID	Sequence	Score	Length	Match	Score	Length	Match
1	095168	15-JUL-1999 (rel. 38, last sequence update)	673	128	99.3%	673	128	99.3%
2	095168	15-JUL-1999 (rel. 38, last sequence update)	509	128	75.1%	509	128	75.1%
3	095168	15-JUL-1999 (rel. 38, last sequence update)	315	133	46.5%	315	133	46.5%
4	095168	15-JUL-1999 (rel. 38, last sequence update)	88.5	1078	13.1%	88.5	1078	13.1%
5	095168	15-JUL-1999 (rel. 38, last sequence update)	71	270	10.5%	71	270	10.5%
6	095168	15-JUL-1999 (rel. 38, last sequence update)	70.5	771	10.4%	70.5	771	10.4%
7	095168	15-JUL-1999 (rel. 38, last sequence update)	70	384	10.3%	70	384	10.3%
8	095168	15-JUL-1999 (rel. 38, last sequence update)	70	695	10.3%	70	695	10.3%
9	095168	15-JUL-1999 (rel. 38, last sequence update)	69.5	294	10.3%	69.5	294	10.3%
10	095168	15-JUL-1999 (rel. 38, last sequence update)	69	298	10.2%	69	298	10.2%
11	095168	15-JUL-1999 (rel. 38, last sequence update)	69	658	10.2%	69	658	10.2%
12	095168	15-JUL-1999 (rel. 38, last sequence update)	69	1023	10.2%	69	1023	10.2%
13	095168	15-JUL-1999 (rel. 38, last sequence update)	69	1660	10.2%	69	1660	10.2%
14	095168	15-JUL-1999 (rel. 38, last sequence update)	68.5	294	10.1%	68.5	294	10.1%
15	095168	15-JUL-1999 (rel. 38, last sequence update)	68.5	574	10.1%	68.5	574	10.1%
16	095168	15-JUL-1999 (rel. 38, last sequence update)	68.5	1333	10.1%	68.5	1333	10.1%
17	095168	15-JUL-1999 (rel. 38, last sequence update)	68	270	10.0%	68	270	10.0%
18	095168	15-JUL-1999 (rel. 38, last sequence update)	67.5	153	10.0%	67.5	153	10.0%
19	095168	15-JUL-1999 (rel. 38, last sequence update)	67	455	9.9%	67	455	9.9%
20	095168	15-JUL-1999 (rel. 38, last sequence update)	67	618	9.9%	67	618	9.9%
21	095168	15-JUL-1999 (rel. 38, last sequence update)	66.5	1036	9.8%	66.5	1036	9.8%
22	095168	15-JUL-1999 (rel. 38, last sequence update)	66	193	9.7%	66	193	9.7%
23	095168	15-JUL-1999 (rel. 38, last sequence update)	66	657	9.7%	66	657	9.7%
24	095168	15-JUL-1999 (rel. 38, last sequence update)	66	677	9.7%	66	677	9.7%
25	095168	15-JUL-1999 (rel. 38, last sequence update)	65.5	450	9.7%	65.5	450	9.7%
26	095168	15-JUL-1999 (rel. 38, last sequence update)	65.5	2358	9.7%	65.5	2358	9.7%
27	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	212	9.5%	64.5	212	9.5%
28	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	315	9.5%	64.5	315	9.5%
29	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	654	9.5%	64.5	654	9.5%
30	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	842	9.5%	64.5	842	9.5%
31	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	1003	9.5%	64.5	1003	9.5%
32	095168	15-JUL-1999 (rel. 38, last sequence update)	64.5	1247	9.5%	64.5	1247	9.5%
33	095168	15-JUL-1999 (rel. 38, last sequence update)	63.5	294	9.4%	63.5	294	9.4%

OY 122 DRTFHLSY 129  
Db 121 DRTFHLSY 128

## RESULT 2

ID NB5M\_BOVIN STANDARD; PRT: 128 AA.

DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE B15 SUBUNIT (EC 1.6.5.3) (EC 1.6.99.3)  
DE (COMPLEX I-B15) (CI-B15).  
GN NDUF84.

OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OA NCBI\_TaxID=9913;  
RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RX TISSUE-Heart;

RA MEDLINE=92389317; PubMed=1518044;

RA Walker J.E., Arizumendi J.M., Dupuis A., Fearnley I.M., Finel M.,

Medd S.M., Pilkington S.J., Runswick M.J., Skehel J.M.,

"Sequences of 20 subunits of NADH:ubiquinone oxidoreductase from

bovine heart mitochondria. Application of a novel strategy for

sequencing proteins using the polymerase chain reaction."

J. Mol. Biol. 226:1051-1072(1992).

CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY

CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED

TO BE UBIQUINONE.

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.

CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC DR EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

CC EMBL: X64898; CAA46107.1;

AC P48306;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE PUTATIVE NADH-UBIQUINONE OXIDOREDUCTASE B15 SUBUNIT (EC 1.6.5.3)  
DE (EC 1.6.99.3) (COMPLEX I-B15) (CI-B15) (HYPOTHETICAL PROTEIN WALTER)  
DE (GGHPW).  
GN NDUF84.

OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OA NCBI\_TaxID=9031;  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94040816; PubMed=7901127;

RA Goldberg G.S., Kaczmarek W.,

"A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1

homeobox is likely to encode the NADH ubiquinone oxidoreductase

subunit B15."

Gene 133:233-235(1993).

CC [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93077061; PubMed=1359990;

RA Goldberg G.S., Kaczmarek W.,

"Sequence of a novel chicken genomic DNA fragment that hybridizes to

the murine Hox-3.1 homeobox."

Gene 121:397-398(1992).

CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY

CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED

TO BE UBIQUINONE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.

CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE

(BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC DR EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

AC P48306;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE PUTATIVE NADH-UBIQUINONE OXIDOREDUCTASE B15 SUBUNIT (EC 1.6.5.3)  
DE (EC 1.6.99.3) (COMPLEX I-B15) (CI-B15) (HYPOTHETICAL PROTEIN WALTER)  
DE (GGHPW).  
GN NDUF84.

OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OA NCBI\_TaxID=9031;  
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94040816; PubMed=7901127;

RA Goldberg G.S., Kaczmarek W.,

"A chicken genomic DNA fragment that hybridizes to the murine Hox-3.1

homeobox is likely to encode the NADH ubiquinone oxidoreductase

subunit B15."

Gene 133:233-235(1993).

CC [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=93077061; PubMed=1359990;

RA Goldberg G.S., Kaczmarek W.,

"Sequence of a novel chicken genomic DNA fragment that hybridizes to

the murine Hox-3.1 homeobox."

Gene 121:397-398(1992).

CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY

CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED

TO BE UBIQUINONE (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.

CC -1- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE

(BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC DR EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.

CC EMBL: X60778; CAA43193.1; ALT-SEQ.



CC Bacteria: Cyanobacteria; Chroococcales; Synechocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
 CC STANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
 CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
 CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
 CC -1- CATALYTIC ACTIVITY: ATP-DEPENDENT BREAKAGE, PASSAGE AND REJOINING  
 CC OF DOUBLE-STRANDED DNA.  
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
 CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
 CC ENZYME FORMS AN A2B2 TETRAMER.  
 CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES  
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)  
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO TYPE II TOPOISOMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D90908; BAA17720.1;  
 DR HSSP: P06982; 1A16.  
 DR InterPro: IPR000565;  
 DR InterPro: IPR001241;  
 DR InterPro: IPR002283;  
 DR InterPro: IPR002288;  
 DR Pfam: PF00986; DNA\_gyraseB\_C; 1.  
 DR Pfam: PF00204; DNA\_topoisomII; 2.  
 DR PRINTS: PR00418; TP12FAMILY.  
 DR PRINTS: PR01159; DNAGYRASEB.  
 DR PROSITE: PS00177; TOPOISOMERASE\_II; 1.  
 DR PROSITE: PS00881; PROTEIN\_SPLICING; FALSE\_NEG.  
 DR K1somerase; Topoisomerase; ATP-binding; Autocatalytic cleavage;  
 K1somerase; Topoisomerase; ATP-binding; Autocatalytic cleavage;  
 K1somerase; Topoisomerase; ATP-binding; Autocatalytic cleavage;  
 FT CHAIN 1 436 GYRB, 1ST PART (POTENTIAL).  
 FT CHAIN 437 871 SSP GYRB INTEIN (POTENTIAL).  
 FT CHAIN 872 1078 GYRB, 2ND PART (POTENTIAL).  
 SQ SEQUENCE 1078 AA; 122819 MW; 6CA02586DFFEA607B CRC64;

Query Match 13.1%; Score 88.5; DB 1; Length 1078;  
 Best Local Similarity 28.2%; Pred. No. 0.24; Mismatches 29; Gaps 4;

Matches 33; Conservative 13; Indels 29; Gaps 4;  
 OY 22 EVNISPEPTRAQAERL-----AIRAQLKREYLLQYNDPNRGLIENPALLRMAYART 73  
 DB 704 EYWAQENNTAQAERVRHFAQNPGLROQYSENAVKQW-----NELLKWRKRT 754  
 OY 74 INVY-PNRPPTPKNSLMGALCGPLFIYIITKTEROKREKLIQCKLDRTHLST 129  
 DB 755 KEQWTEPFREKREALAQT-----YRKTLAALKQVEIENGYLIDISADSY 800

RESULT 5  
 FRG\_ZYMO  
 ID FRG\_ZYMO STANDARD: PRT: 270 AA.  
 AC Q9X3X1;

DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (EC 3.2.2.23) (FAPY-DNA  
 DE GLYCOSYLASE).  
 GN MOTT OR FPG.  
 OS Zymomonas mobilis.  
 OS Bacteria: Proteobacteria; alpha subdivision; Sphingomonas group;  
 OC Zymomonas.  
 OX NCBI\_TaxID=542;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 31821 / 2M4 / CP4;  
 RA Lee H.J., Kang H.S.;  
 RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THIS ENZYME MAY PLAY A SIGNIFICANT ROLE IN PROCESSES  
 CC LEADING TO RECOVERY FROM MUTAGENESIS AND/OR CELL DEATH BY  
 CC ALKYLATING AGENTS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE DEOXYRIBOSE N-GLYCOSIDIC  
 CC BOND TO EXCISE 2,6-DIAMINO-4-HYDROXY-5N-METHYLFORMAMIDOPYRIMIDINE  
 CC (FAPY) OR 4,6-DIAMINO-5-FORMAMIDOPYRIMIDINE.  
 CC -1- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE FPG FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF088896; AAD21547.1; ALT\_INT.  
 DR InterPro: IPR000191;  
 DR Pfam: PF01149; Fapy\_DNA\_glyco; 1.  
 DR PROSITE: PS01242; FPG; 1.  
 DR DNA repair; Hydrolyase; Glycosidase; Zinc.  
 FT ZN FING 246 268 POTENTIAL.  
 SQ SEQUENCE 270 AA; 30693 MW; EED0E1943DF5F58 CRC64;

Query Match 10.5%; Score 71; DB 1; Length 270;  
 Best Local Similarity 36.2%; Pred. No. 3.1;  
 Matches 17; Conservative 6; Mismatches 18; Indels 6; Gaps 2;

OY 42 QLRREYLLQYNDPNRKG---LIENPALLRMAYARTIVP---NFRP 82  
 DB 96 QTNVNFIVSLYDPRFGSLDLVKNQLEWSYFRNIGPEPLTGTFNP 142

RESULT 6  
 DNK3\_SYNY3  
 ID DNK3\_SYNY3 STANDARD: PRT: 771 AA.  
 AC P73098;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE DNAK PROTEIN 3 (HEAT SHOCK PROTEIN 70) (HSP70).  
 GN DNAK3 OR SLI1932.  
 OS Synechocystis sp. (strain PCC 6803).  
 OS Bacteria: Cyanobacteria; Chroococcales; Synechocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";

RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).  
 CC -1- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D90903; BAA17123.1; -  
 DR HSSP: P19120; IATS.  
 DR InterPro: IPR001023; -  
 DR Pfam: PF00012; HSP70.1.  
 DR PRINTS: PR00301; HEATSHOCK70.  
 DR PROSITE: PS00297; HSP70.1; 1.  
 DR PROSITE: PS00329; HSP70.2; 1.  
 DR PROSITE: PS01036; HSP70.3; FALSE\_NEG.  
 DR Rosette: ATP-binding; Multigene family; Heat shock.  
 KW Chaperone: ATP-binding; Multigene family; Heat shock.  
 SO SEQUENCE 771 AA; 86030 MW; BE715F28B372738C CRC64;  
  
 Query Match 10.4%; Score 70.5; DB 1; Length 771;  
 Best Local Similarity 30.5%; Pred. No. 12;  
 Matches 25; Conservative 9; Mismatches 33; Indels 15; Gaps 3;  
  
 Oy 16 EFTDPAEYNISPTERRAQAERLAIRAOIKREYLYQNDPNRRGLIENPALLRMAYARTI-74  
 Db 570 EILDSLEKDERLDAQAQADLVLELNREVRQYDD-KEEGFE-----AIKKTFT 621  
 Oy 75 -----NYPNRPPTPKNSLNG 90  
 Db 622 GDFDDDDYNNRRPAPRDYRG 643  
  
 RESULT 7  
 SX18\_HUMAN STANDARD; PRT; 384 AA.  
 AC P35713; Q9NPH8;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSCRIPTION FACTOR SOX-18.  
 GN SOX18.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed=10858556;  
 RA Stanojic S., Stevanovic M.;  
 RT "The human SOX18 gene: cDNA cloning and high resolution mapping";  
 RL Biochim. Biophys. Acta 1492:237-241(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Azuma T., Seki N., Yoshikawa T., Masuho Y., Muramatsu M.;  
 RT "cDNA cloning, tissue expression and chromosome mapping of human  
 RT homolog of SOX18";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 96-149 FROM N.A.  
 RX MEDLINE=92310993; PubMed=1614875;  
 RA Denny P., Swift S., Brand N., Dabhade N., Barton P., Ashworth A.;  
 RT "A conserved family of genes related to the testis determining gene,  
 RT SRY";  
 CC Nucleic Acids Res. 20:2887-2887(1992).  
 CC -1- FUNCTION: BINDS TO THE CONSENSUS SEQUENCE 5'-AACAAAG-3' AND IS  
 CC ABLE TO TRANS-ACTIVATE TRANSCRIPTION VIA THIS SITE (BY

CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
 CC -1- SIMILARITY: CONTAINS 1 HMG BOX.  
 CC -1- CAUTION: WAS CALLED SOX-8 BY REF.3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ243896; CAB95835.1; -  
 DR EMBL: AB033888; BAA94874.1; -  
 DR EMBL: X65664; CAA46615.1; -  
 DR PIR: S21482; S21482.  
 DR PIR: S22941; S22941.  
 DR HSSP: Q05066; IHR2.  
 DR MIM: 601618; -  
 DR InterPro: IPR000910; -  
 DR Pfam: PR00505; HMG\_box; 1.  
 KW Transcription regulation; Activator; DNA-binding; Nuclear protein.  
 FT DNA\_BIND 85 153 HMG BOX.  
 SO SEQUENCE 384 AA; 40891 MW; 327462E519770062 CRC64;  
  
 Query Match 10.3%; Score 70; DB 1; Length 384;  
 Best Local Similarity 20.8%; Pred. No. 6;  
 Matches 31; Conservative 21; Mismatches 43; Indels 54; Gaps 5;  
  
 Oy 4 PKYPSRLTLETLDAEYNISPTERRAQAERLAIR-----AOLKREYLYQYN 52  
 Db 48 PASPPSPQSPRSPSPERGVLGSPAGRGROADESRIRPMNAFMWAKDERRLAQN 107  
 Oy 53 D-----PNRRGLIENPALLRMAYARTINYPNRPPTPKNSLNGAL 92  
 Db 108 PDLHNAVLSKMLGKAMKELNAEKRPFEAEERLRYOHLRD--HPNRYR----- 157  
 Oy 93 CGFGLIFITYIITKTERDKELIOEGL 121  
 Db 158 -----KKQARKARKLEPGL 173  
  
 RESULT 8  
 VATI\_MENJA STANDARD; PRT; 695 AA.  
 ID VARI\_MENJA  
 AC Q57675;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE V-TYPE ATP SYNTHASE SUBUNIT I (Ec 3.6.1.34) (V-TYPE ATPASE SUBUNIT I).  
 GN ATP1 OR M10222.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sult G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utechtack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii";  
 RL Science 273:1058-1073(1996).  
 CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON

CC GRADIENT ACROSS THE MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U67478; AAB98208.1; -  
CC TIGR: MJ0222; -  
CC DR InterPro: IPR002490; -  
CC Pfam: PF01496; V.Arpaase-sub.a; 1.  
CC KMW Hydroxylase; Hydrogen ion transport; Transmembrane.  
CC FT TRANSMEM 392 412 POTENTIAL.  
CC FT TRANSMEM 425 445 POTENTIAL.  
CC FT TRANSMEM 497 517 POTENTIAL.  
CC FT TRANSMEM 534 554 POTENTIAL.  
CC FT TRANSMEM 556 576 POTENTIAL.  
CC FT TRANSMEM 605 625 POTENTIAL.  
CC FT TRANSMEM 627 647 POTENTIAL.  
CC SQ SEQUENCE 695 AA; 76953 MW; F03E5CEBEE29D53D CRC64;  
  
Query Match 10.3%; Score 70; DB 1; Length 695;  
Best Local Similarity 23.6%; Pred. No. 12;  
Matches 29; Conservative 17; Mismatches 41; Indels 36; Gaps 3;  
  
QY 13 TLPLETDAEYNISPEPTRAQAERLAI-----RAQLKREYLLOYNDPNRGLIENPA 64  
DB 217 TLKEKLE-----NVLSIKKFEFEYRDISDECTPSEALSKISELSERNSLIEKLK 272  
QY 65 LLRAAYATINYPFRFPKNSLGCALGCPPLFIYIITERTDRKKLLQEGLDRT 124  
DB 273 ALAQKWER-----ELLAVALLELSIKRAGDAYSQFGKTDRT 308  
  
QY 125 FHL 127  
DB 309 YTI 311  
  
RESULT 9  
CC21\_ORYSA STANDARD; PRT; 294 AA.  
AC P29618;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG 1 (EC 2.7.1.-).  
OS CDC2-1.  
GN Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;  
OC Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV\_NIPONBARE;  
RX MEDLINE=9293101; PubMed=1376401;  
RA Hasegawa J., Hirabayashi T., Hayano Y., Hata S., Ohashi Y.,  
RA Suzuki I., Utsugi T., Toh-E A., Kikuchi Y.;  
RT Isolation and characterization of cDNA clones encoding cdc2  
RT homologues from Oryza sativa: a functional homologue and cognate  
RT variants.\*.  
RL Mol. Gen. Genet. 233:10-16(1992).  
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC  
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES  
CC THE REPLICATIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.  
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES  
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY  
CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC CDC2/CDCX SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X60374; CAA42922.1; -  
CC DR PIR: S22440; S22440.  
CC DR HSP: P24941; 1HKX.  
CC DR InterPro: IPR000719; -  
CC DR InterPro: IPR002290; -  
CC Pfam: PF00069; pkinase; 1.  
CC DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
CC DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC KMW Transferase; Serine/threonine-protein kinase; ATP-binding;  
CC cell cycle; Cell division; Mitosis; Phosphorylation.  
CC FT DOMAIN 4 287 PROTEIN KINASE.  
CC FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
CC FT BINDING 33 33 ATP (BY SIMILARITY).  
CC FT ACT\_SITE 127 127 BY SIMILARITY.  
CC FT MOD\_RES 14 14 PHOSPHORYLATION (BY SIMILARITY).  
CC FT MOD\_RES 15 15 PHOSPHORYLATION (BY SIMILARITY).  
CC FT MOD\_RES 161 161 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
CC SQ SEQUENCE 294 AA; 34071 MW; 51322D93AEF4C131 CRC64;  
  
Query Match 10.3%; Score 69.5; DB 1; Length 294;  
Best Local Similarity 36.2%; Pred. No. 5;  
Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;  
  
QY 2 SFPRKYPSSLKTLPLETDAEYN-ISPEPTRAQAERLAIQAOLKREY 47  
DB 240 AFRKQAOADLPIVPLDPAAGDLLSKMLKYEPPNKRITARQALEHEX 286  
  
RESULT 10  
VPO\_BPHP1 STANDARD; PRT; 298 AA.  
ID VPO\_BPHP1  
AC P51719;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE PROBABLE CAPSID SCAFFOLDING PROTEIN (ORF17).  
OS Bacteriophage HPI.  
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.  
OX NCBI\_TaxID=10650;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HPI1;  
RX MEDLINE=96279738; PubMed=8710508;  
RA Esposito D., Filizmaurice W.P., Benjamin R.C., Goodman S.D.,  
RA Waldman A.S., Socca J.J.;  
RT "The complete nucleotide sequence of bacteriophage HPI DNA".;  
RL Nucleic Acids Res. 24:2360-2368(1996).  
CC -1- SIMILARITY: STRONG, TO PHAGE P2 PROTEIN O.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U24159; AAB09202.1; -  
CC Capsid assembly.  
CC SQ SEQUENCE 298 AA; 33702 MW; 6D284146BD700C1F CRC64;





```

Db 790 SEQYGPGENULPET--DANKFSEDEEQOQOLREEQOQEREOLOLEREQOER 847
OY 53 DPNRGULE-IPALLRMVARTINYPNRP-----TPKNSLMGALCGFLIFTIYIIR 106
Db 848 EOOEGGQOEPGP-----EEYPSYEYSRALOEKMNMRDRI-----YAE 886
OY 107 TERDR--KEKLOEGK 120
Db 887 QERERQOETILOENO 902

RESULT 13
VIT6_OSCBR STANDARD: PRT: 1660 AA.
AC 094637:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE VITELLOGENIN 6 PRECURSOR.
GN VIT-6.
OS Oscheilus brevesophaga.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Rhabditinae; Oscheilus.
OX NCBI_TaxID=57871;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CEM1:
RX MEDLINE=96212989; PubMed=8676742;
RA Winter C.E., Penha C., Blumenthal T.;
RT "Comparison of a vitellogenin gene between two distantly related
RT rhabditid nematode species.";
RL Mol. Biol. Evol. 13:674-684(1996).
-1- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE SOURCES OF
NUTRIENTS DURING EMBRYONIC DEVELOPMENT (POTENTIAL).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U35449; AAB49749.1; -
DR HSSP: P11064; IPHR.
DR InterPro: IPR001747; -
DR InterPro: IPR001846; -
DR fam: PRO1347; Vitellogenin_N; 1.
KW Storage protein; Multigene family; Signal.
FT CHAIN 1 1660
FT SIGNAL 15
FT CARBOHYD 16 1660
FT CARBOHYD 237 383
FT CARBOHYD 693 693
FT CARBOHYD 1307 1307
FT CARBOHYD 1596 1596
FT CARBOHYD 1629 1629
FT SEQUENCE 1660 AA; 192108 MW; 2F78263C0678ABBD CMC64;

Query Match 10.2%; Score 69; DB 1; Length 1660;
Best local similarity 24.8%; Pred. No. 43;
Matches 34; Conservative 20; Mismatches 57; Indels 26; Gaps 4;
OY 6 YKRSRLTLPETIDPPEYNISPTERRAQARLAKROKREYLYQVDPNRRGLIENPAL 65
Db 631 YTWSTLTKTJSESENPAE---KEIRRVQSOSLASIPEEQKYLESKHTNMFNMGSCAT 686
OY 66 LKMAVARTINYPNRPNSLMGALCGF-----PLIFTI-----YIITE 108
Db 687 LKMA-----TIFSNDSVLKREITASLETYFGGEMKNKYLAOIGLYQNNLDSVLKLOKVE 741

```

```

OY 109 RDRKEKLOEGKLDRTF 125
Db 742 ETGLEQLVVRGKRSSSF 758

RESULT 14
CC2A_ARATH STANDARD: PRT: 294 AA.
ID CC2A_ARATH
AC P24100;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 2 HOMOLOG A (EC 2.7.1.-).
GN CDC2A OR CDC2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93005715; PubMed=1840925;
RA Ferreira P.C.G., Hemery A.S., Villarroel R., van Montagu M., Inze D.;
RT "The Arabidopsis functional homolog of the p34cdc2 protein kinase.";
RL Plant Cell 3:531-540(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA.
RX MEDLINE=92039027; PubMed=1937013;
RA Hirayama T., Imajuku Y., Anal T., Matsui M., Oka A.;
RT "Identification of two cell-cycle-controlling cdc2 gene homologs in
RT Arabidopsis thaliana.";
RL Gene 105:159-165(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92316202; PubMed=1618302;
RA Imajuku Y., Hirayama T., Endoh H., Oka A.;
RT "Exon-intron organization of the Arabidopsis thaliana protein kinase
RT genes CDC2a and CDC2b.";
RL FEBS Lett. 304:73-77(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339744; PubMed=1634002;
RA Inze D., Ferreira P.C.G., Hemery A.S., van Montagu M.;
RT "Control of cell division in plants.";
RL Biochem. Soc. Trans. 20:80-84(1992).
CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC
CC CELL CYCLE. COMPONENT OF THE KINASE COMPLEX THAT PHOSPHORYLATES
CC THE REPETITIVE CARBOXYL-TERMINUS OF RNA POLYMERASE II.
CC -1- ENZYME REGULATION: PHOSPHORYLATION AT THR-14 OR TYR-15 INACTIVATES
CC THE ENZYME, WHILE PHOSPHORYLATION AT THR-161 ACTIVATES IT (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC CDC2/CDCX SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: S45387; AAB32643.1; -
DR EMBL: M59198; AAA32831.1; -
DR EMBL: D10850; BAA01623.1; -
DR EMBL: X57839; CAA40971.1; -
DR PIR: A48984; A48984.
DR PIR: J00967; J00967.
DR PIR: J01337; J01337.
DR PIR: S23095; S23095.
DR HSSP: P24941; IHCK.

```

DR InterPro: IPR000719; -  
 DR InterPro: IPR002290; -  
 DR Pfam: PF00069; PKinase; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 4 287  
 FT NP\_BIND 10 18  
 FT BINDING 33 33  
 FT ACT\_SITE 127 127  
 FT MOD\_RES 14 14  
 FT MOD\_RES 15 15  
 FT MOD\_RES 161 161  
 SO SEQUENCE 294 AA; 34030 MW; B5FAE5FA9EC366E CRC64;

Query Match 10.1%; Score 68.5; DB 1; Length 294;  
 Best Local Similarity 36.2%; Pred. No. 6.3;  
 Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;  
 OY 2 SEPKYKPSLRLPETLPAEYN-ISPETRRQARLAIKRLKREY 47  
 DB 240 AFPKWKPTDLEFVNPDPDGVLLSKMLMDPTKRINARAALREHY 286

RESULT 15  
 YJ23\_YEAST  
 ID YJ23\_YEAST STANDARD: PRT; 574 AA.  
 AC P47113;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHEETICAL. 66.1 KDA PROTEIN IN RAD7-HITI INTERGENIC REGION.  
 GN YJ0053W OR J1667.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C;  
 RX MEDLINE=95397595; PubMed=7668047;  
 RA Huang M.-E., Chuat J.-C., Galibert F.;  
 RT "Analysis of a 42.5 kb DNA sequence of chromosome X reveals three  
 RT tRNA genes and 14 new open reading frames including a gene most  
 RT probably belonging to the family of ubiquitin-protein ligases.";  
 RT Yeast 11:775-781(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: L36344; AAA8756.1; -  
 DR EMBL: Z49553; CA89581.1; -  
 KW Hypothetical protein.  
 FT DOMAIN 193 201  
 FT DOMAIN 265 268  
 FT POLY-SER.  
 SO SEQUENCE 574 AA; 66086 MW; 1EA75D26FCF0A1DA CRC64;

Query Match 10.1%; Score 68.5; DB 1; Length 574;  
 Best Local Similarity 22.4%; Pred. No. 14;  
 Matches 22; Conservative 20; Mismatches 23; Indels 33; Gaps 5;

OY 4 PKYKSSSLRLPETLPAEYN-----SPTERRAQ-----AERLAIR----- 40  
 DB 259 PFLFKSSSSSLPKLISPAQYDIVKHEDELLTGLHRRQRDMNTQOELDSFKERKSVRHCSN 318

OY 41 -----AQLKR-EYLLQYNDPNRRG-LIENPALLRW 68  
 DB 319 QNVQLNGPAKIKITIKQIDHNTPMKKGSMILYNKTKMK 356

Search completed: April 29, 2001, 06:47:26  
 Job time: 132 sec







RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pachet J.M.,  
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang M., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003813; AAF58210.1; -  
 DR FLVBASE: FBgn0033961; CG12859.  
 SQ SEQUENCE 113 AA; 13134 MW; 25A8CDB90F2DB414 CRC64;

Query Match 14.2%; Score 96.5; DB 5; Length 113;  
 Local Similarity 29.3%; Pred. No. 0.01; Mismatches 36; Indels 9; Gaps 3;  
 Matches 27; Conservative 20;

OY 42 QLRKREYLQYNDPNRR-----GLIENPALLRMAYARTINVPNRPPTKSLMGALCGFG 96  
 DB 20 KLRQELKSSNPYRIATGEGTVPDAGLARFGAMRVSN-YEHFKPTGKFRIGLFAVVL 78  
 OY 97 PLFIYIYIKTERDRREKLIQESKL--DRTF 125  
 DB 79 PIALYAMALKAERDGRREKYRTGOVAVKDRDF 110

RESULT 2  
 O9L7P6 PRELIMINARY; PRT; 386 AA.  
 AC O9L7P6;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE PARC (FRAGMENT).  
 GN PARC.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_Taxid=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-KIM;  
 RA Lindler L.E., Jahan N.;  
 RT Non-radioactive detection of ciprofloxacin resistance in *Yersinia*  
 RL pestis.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF221695; AAF36389.1; -  
 FT NON\_TER 1  
 FT NON\_TER 386  
 SQ SEQUENCE 386 AA; 43477 MW; CC748E6024664CD0 CRC64;

Query Match 12.1%; Score 82; DB 2; Length 386;  
 Best Local Similarity 27.7%; Pred. No. 1.5; Mismatches 59; Indels 18; Gaps 5;  
 Matches 36; Conservative 17;

OY 11 LRTLPETLDAEYNISPEERRAQAER-----LAIRAQLKREYLQY-----NDPNRGLI 60  
 DB 252 LRDESCHENPTRLIVIPRTNRVLDQVMSHLFTTDERSYRINMNMIGLDNRPVAGL 311  
 OY 61 ENPALLRMAYAR--TINVYNFRPTP-----KNSLMGALCGFGPLFIYIYIKTERDRKREK 114  
 DB 312 E--ILTEWLVEFRQTVNRLNRLERLEKVLKRLHILEGLLFLNIDEVIIHITREDDKPL 369

OY- 115 LIOEGKLDRT 124  
 DB 370 LMKRESISET 379

RESULT 3  
 O9K198 PRELIMINARY; PRT; 313 AA.  
 AC O9K198;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE HYPOHERETICAL 37.1 KDA PROTEIN.  
 OS Bacteroides uniformis.  
 OC Bacteria; CFN group; Bacteroidaceae; Bacteroides.  
 OX NCBI\_Taxid=820;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20115334; PubMed=10648516;  
 RA Shoemaker N.B., Wang G.R., Salyers A.A.;  
 RT "Multiple gene products and sequences required for excision of the  
 RT mobilizable integrated Bacteroides element NB01.";  
 RL J. Bacteriol. 182:928-936(2000).  
 DR EMBL: AP238307; AAF74439.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 313 AA; 37128 MW; DID43577FDB882D9 CRC64;

Query Match 11.0%; Score 74.5; DB 2; Length 313;  
 Best Local Similarity 22.4%; Pred. No. 6.9; Mismatches 19; Indels 11; Gaps 3;  
 Matches 15; Conservative 22;

OY 4 PKYPSRLTLPETLDAEYNISPEERRAQAERLAIRAQLKREYLQYNDPNRGLIENP 63  
 DB 51 PKYEAARKRIGDDRRNP-DFE-----ELEENVLOSLEAKIYNEHTPNKRSIV--P 99  
 OY 64 ALRMAY 70  
 DB 100 AMLAMLY 106

RESULT 4  
 O9WUN8 PRELIMINARY; PRT; 195 AA.  
 AC O9WUN8;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE LADYBIRD-LIKE HOMODOMAIN PROTEIN Lbx2.  
 GN Lbx2H OR Lbx2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chen F., Liu K.C., Epstein J.A.;  
 RT "Lbx2, a novel murine homobox gene related to the *Drosophila* ladybird  
 RT gene is expressed in the developing urogenital system, eye, and  
 RT brain.";  
 RL Mech. Dev. 0:0-0(1999).  
 DR EMBL: AF146150; AAD31905.1; -  
 DR HSSP: P23441; 1FTT.  
 DR MGD: MGI:1342288; Lbx2h.  
 DR INTERPRO: IPR000047; -  
 DR INTERPRO: IPR001356; -  
 DR PFAM: PF00046; homobox; 1.  
 DR PRINTS: PR00024; HOMEOBOX.  
 DR PRINTS: PR00031; HTHREPRESSR.  
 DR PROSITE: PS50071; HOMEOBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein.  
 SQ SEQUENCE 195 AA; 20916 MW; E106CE108933D845 CRC64;

Query Match 10.9%; Score 74; DB 11; Length 195;  
 Best Local Similarity 29.2%; Pred. No. 4.4;

Matches 33; Conservative 15; Mismatches 39; Indels 26; Gaps 7;

Query Match 10.9%; Score 74; DB 1; Length 451;  
 Best Local Similarity 32.6%; Pred. No. 12;  
 Matches 29; Conservative 10; Mismatches 26; Indels 24; Gaps 5;

QY 11 LRLPPELDPAEVINSPT-----RRAOERLAIRAOIKREYLYQNDPNRGLIE 61  
 DB 29 LPDPELPDP-----LDPEEPIDIEKLRIFAEL-VKQETSR-----RY-----IE 73

QY 62 NPALLRMAYARTINVPNFRPTPKNSL 90  
 DB 74 IPGELRLKLYSKIGRPPLFRATNLEKLG 102

RESULT 6  
 ID 09WXH9 PRELIMINARY; PRT: 526 AA.  
 AC 09WXH9;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)  
 OS Lactobacillus casei.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Lactobacillus.  
 NCBI\_TaxID=1582;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 27092;  
 RA Senba M., Nakashima Y., Miake F., Watanabe K.;  
 RT Cloning and Expression of the beta-N-Acetylglucosaminidase Gene from  
 RT Lactobacillus casei ATCC 27092 and Characteristics of the enzyme  
 RT expressed in *Escherichia coli*.  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB025100; BAA76352.1;  
 SQ SEQUENCE 526 AA; 56835 MW; 255D541FA300C723 CRC64;

Query Match 10.9%; Score 74; DB 2; Length 526;  
 Best Local Similarity 30.0%; Pred. No. 15;  
 Matches 27; Conservative 17; Mismatches 36; Indels 10; Gaps 4;

QY 10 SLRTLPETLDPAEVINSPTERRAOERLAIRAOIKREYLYQNDPNRGLIEN--PALLR 67  
 DB 60 ALRL--LDDMSVINGTSGSDVVKQALQGNK-----AYDDPFGNALTQTDMDAVLK 112

QY 68 WARTINVPNFR-PTPKNSLGMALCGRG 96  
 DB 113 YAAARDINIPIVNSPGHMDALITAMAQLG 142

RESULT 7  
 ID 09LFZ1 PRELIMINARY; PRT: 520 AA.  
 AC 09LFZ1;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,  
 RA Shinn P., Altafi H., Bel Q., Chin C., Chlou J., Choi E., Conn L.,  
 RA Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Tortum M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F20N2 from chromosome  
 RT 1."  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Buehler E.E., Dunn P.P., Feng J.J., Kim C.C., Li Y.Y.,  
 RA Walker M.M., Altafi H., Araujo R.R., Conn L.L., Conway A.A.B.,  
 RA Gonzalez A.A., Hansen N.N.F., Huizar L.L., Kremenetska I.I.,  
 RA Lenz C.C., Li J.J., Liu S.S., Lueros S.S., Rowley D.D., Schwartz J.J.,  
 RA Tortum M.M., Vysotska V.V., Yu G.G., Davis R.R.W.,  
 RA Federspiel N.N.A., Theologis A.A., Ecker J.J.R.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bel Q., Chin C., Chlou J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharzky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Tortum M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC002328; AAF79499.1;  
 SQ SEQUENCE 520 AA; 60115 MW; 3945B20167DCC56B CRC64;

Query Match	10.4%	Score 70.5;	DB 10;	Length 585;
-best Local Similarity	26.9%	Pred. No. 39;		
Matches	28;	Conservative	18;	Mismatches 39;
				Indels 19;
				Gaps 3;

NCBI_TaxID=9606;	OX
[1]	RN

RA Stanojic S., Stevanovic M.;  
 RT "The human SOX18 gene: cDNA cloning and high resolution mapping."  
 RL Blochm. Biophys. Acta 1492:237-241(2000).  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Azuma T., Seki N., Yoshikawa T., Masuho Y., Muramatsu M.;  
 RT "cDNA cloning, tissue expression and chromosome mapping of human  
 RT homolog of SOX18."  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN 13  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Dunn T.L., Myneett-Johnson L., Wright E.M., Hosking B.M., Koopman P.A.,  
 RA Muscat G.E.;  
 RT "Sequence and expression of sox-18 encoding a new HMG-box  
 RT transcription factor."  
 RL Gene 161:223-225(1995)  
 DR EMBL; AJ243886; CAB95835.1;  
 DR EMBL; AB033888; BAA94874.1;  
 SQ SEQUENCE 384 AA; 40891 MW; 327462519770062 CRC64;

Query Match 10.3%; Score 70; DB 4; Length 384;  
 Best Local Similarity 20.8%; Pred. NO. 26;  
 Matches 31; Conservative 21; Mismatches 43; Indels 54; Gaps 5;

OY 4 PKYPSRLRTPELIDPAEYVNISPETRAQAEKRLAIR-----AOLKREYLQYN 52  
 DB 48 PASPPSPQSRSPRSPERPGYGLSPAGRGEOAADSRIKRRPNATWAKDKRKLQON 107  
 OY 53 D-----PNRGLIENPALLMAYARTINVPNFRPPKNSLMGAL 92  
 DB 108 PDLHNAVLKMLGAKMKELNAEKRPFEVEAEKRLVQLRD---HPNKKYRBR----- 157  
 OY 93 CGFGPLFIYYIKTERDKREKLEOGKL 121  
 DB 158 -----RKKQARKARRLEPGLL 173

RESULT 11  
 Q25693 PRELIMINARY; PRT; 1025 AA.  
 AC Q25693;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE PWDNR2 PROTEIN.  
 CN PWDNR2.  
 OC Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-D10;  
 RX MEDLINE=94333528; PubMed=7914495;  
 RA "Radio J.P., Cowman A.F.;  
 RT "Plasmodium falciparum: the pfmdr2 protein is not overexpressed in  
 RT chloroquine-resistant isolates of the malaria parasite."  
 RL Exp. Parasitol. 79:137-147(1994).  
 DR EMBL; U04640; AAA21513.1;  
 DR HSSP; P13569; 1MBD.  
 DR INTERPRO; IPR001140;  
 DR INTERPRO; IPR001617;  
 DR PFWAM; PF00005; ABC\_tran. 1.  
 DR PFWAM; PF00664; ABC\_membrane. 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER. 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1025 AA; 119152 MW; 91AF3BD8916D520D CRC64;

Query Match 10.3%; Score 70; DB 5; Length 1025;  
 Best Local Similarity 33.3%; Pred. NO. 86;

Matches 14; Conservative 11; Mismatches 17; Indels 0; Gaps 0;  
 OY 85 KNSMGALCGFGLFIYYIKTERDKREKLEOGKLDRTFH 126  
 DB 500 KNSLGSVLEFGLITLYSTIKTKRKRRKIRKANEMDNVNH 541

RESULT 12  
 Q9XF46 PRELIMINARY; PRT; 280 AA.  
 AC Q9XF46;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE CELL DIVISION CONTROL PROTEIN 2 (FRAGMENT).  
 OS Phaseolus aureus (Mung bean) (Vigna radiata).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Vigna.  
 OX NCBI\_TaxID=3916;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Luo S., Liu S.-L., Wang Y., Liang H., Wang Y.-C., Han B.-W.;  
 RT "Cdc2 gene cloning and its expression along with differentiation of  
 RT adventitious root in mung bean."  
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF129886; AAD30506.1;  
 DR HSSP; P24941; 1HC.  
 DR INTERPRO; IPR000719;  
 DR INTERPRO; IPR001245;  
 DR INTERPRO; IPR002290;  
 DR PFWAM; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST. 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM. 1.  
 KW Cell division.  
 FT NON\_TER  
 SQ SEQUENCE 280 AA; 32166 MW; 85826C20A864253E CRC64;

Query Match 10.3%; Score 69.5; DB 10; Length 280;  
 Best Local Similarity 34.0%; Pred. NO. 20;  
 Matches 16; Conservative 12; Mismatches 18; Indels 1; Gaps 1;  
 OY 2 SPFKYPSRLRTPELIDPAEYVNISPETRAQAEKRLAIRAQLKREY 47  
 DB 230 APPKMPKDLKTYVPLDPAAGLDLSRLMLDPSKRITGRSALEHEV 276

RESULT 13  
 Q9XF13 PRELIMINARY; PRT; 280 AA.  
 AC Q9XF13;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE CELL DIVISION CONTROL PROTEIN 2 (FRAGMENT).  
 OS Phaseolus vulgaris (Kidney bean) (French bean).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Phaseolus.  
 OX NCBI\_TaxID=3885;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Luo S., Liu S.-L., Wang Y., Wang Y.-C., Han B.-W.;  
 RT "Cdc2 gene cloning and its expression along with differentiation of  
 RT adventitious root in mung bean."  
 RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF126737; AAD30494.1;  
 DR INTERPRO; IPR000719;  
 DR INTERPRO; IPR001245;  
 DR INTERPRO; IPR002290;  
 DR PFWAM; PF00069; pkinase; 1.

DR PRINTS; PR00109; TYRKINASE.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Cell division.  
FT NON\_TER 1  
SEQUENCE 280 AA; 32166 MW; 85826C20A864253E CRC64;

Query Match 10.3%; Score 69.5; DB 10; Length 280;  
Best Local Similarity 34.0%; Pred. No. 20;  
Matches 16; Conservative 12; Mismatches 18; Indels 1; Gaps 1;

OY 2 SFEYKPKSSLRLPETLDAPEYN-ISPETRAQAERLAIRAQLKREY 47  
:111:1 1:1 111: :1 :1:1:1:1  
DB 230 AFPMQPKDLKTVVNPDLPAGLDLRLHLDPKRTIGRSALHEHY 276

RESULT 14  
ID 092B47 PRELIMINARY; PRT; 742 AA.  
AC 092B47;  
DT 01-MAY-1999 (TREMBLrel. 10; Created)  
DT 01-MAY-1999 (TREMBLrel. 10; Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13; Last annotation update)  
DE CPA.  
GN Streptococcus pyogenes.  
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS101;  
RA Podbielski A., Wolschik M., Leonard B.A.B., Schmidt K.H.;  
RT "Characterization of nra, a global negative regulator gene in group A streptococci."  
RL Mol. Microbiol. 31:0-0(1999).  
DR EMBL: U49397; AAC97148.1; -;  
SO SEQUENCE 742 AA; 83682 MW; 97A1FF44BAECB944 CRC64;

Query Match 10.3%; Score 69.5; DB 2; Length 742;  
Best Local Similarity 24.8%; Pred. No. 65;  
Matches 27; Conservative 17; Mismatches 44; Indels 21; Gaps 5;

OY 8 PSSLRTPETLDAPEYNISPETRAQAERLAIRAQLKREYLQYNDPNRGLIENPALLR 67  
DE 319 PNGTYTLTETSSPDGYK-----AEPKFRVENKKVFIYQKDSQ---VENPN-KE 365  
OY 68 WAYARTINYPNRPPTPKNSLMGALCGFGLIFITYIKTERDRKREKLI 116  
DB 366 VAEPYVEAYNDPMDE-----VLSGFYPGKFY--ATNKDKSSQV 406

RESULT 15  
ID 0918D1 PRELIMINARY; PRT; 1276 AA.  
AC 0918D1;  
DT 01-OCT-2000 (TREMBLrel. 15; Created)  
DT 01-OCT-2000 (TREMBLrel. 15; Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15; Last annotation update)  
DE MYOSIN VI.  
GN CMV6.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kendrick-Jones J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
RA Buss F., Kendrick-Jones J., Lionne C., Knight A.E., Cote G.P.,  
RA Luzzio P.;  
RT "The localization of myosin VI at the golgi complex and leading edge  
RT of fibroblasts and its phosphorylation and recruitment into membrane  
RT ruffles of A431 cells after growth factor stimulation."  
RL J. Cell Biol. 143:1535-1545(1998).  
DR EMBL: A1278608; CAB96536.1; -;  
SO SEQUENCE 1276 AA; 147603 MW; 5F6AEE43A6FB260F CRC64;

Query Match 10.3%; Score 69.5; DB 13; Length 1276;  
Best Local Similarity 21.7%; Pred. No. 1.3e+02;  
Matches 33; Conservative 30; Mismatches 46; Indels 43; Gaps 8;

OY 5 KYKPSLRPLPETLDAPEYNISPETRR-----QAERLAIRAQL-----KREYLQY-- 51  
:11:11111 :1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1  
DB 114 KYGGRSLGTLF---PHVAIADKAYRDMKVKLMSSQIVSGSGAGKENTFVLRILT 169  
OY 52 -----NDPNRGLIENPALLRWAVARTI-----NYPNRPPTPKNSLMGALCGFGP 97  
DB 170 ESTGTGDDIDRIVEANPLLEAFGNKAKTINNNSRGRKVEIHFNKKNVVG---GF-- 224  
OY 98 LFIYIYIKTERDRKREKLIOEGRKLDRTFHLISY 129  
:1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1 :1:1:1:1  
DB 225 --VSHYLL-----KSRICVGKKEERNYHIFY 249

Search completed: April 29, 2001, 06:47:58  
Job time: 139 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 29, 2001, 06:34:44 ; Search time 20.28 Seconds

(Without alignments)  
363.612 Million cell updates/sec

Title: US-09-726-899-3

Perfect score: 678

Sequence: 1 MSFPRKPSRLTLPETLDP.....DRKEKIQGKLDRTFLSY 129

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

T number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
A.Geneseq\_0401:\*  
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	19	W69225
2	503	74.2	113	20	Y76629
3	76	11.2	16	21	B53504
4	72	10.6	290	21	G33365
5	69.5	10.3	292	21	B35797
6	69.5	10.3	294	20	W95690
7	69	10.2	656	17	R88358
8	69	10.2	660	17	R88359
9	69	10.2	660	21	Y96186
10	68.5	10.1	204	21	G34044
11	68.5	10.1	237	21	G11215

12	68.5	10.1	237	21	G54043
13	68.5	10.1	294	21	G11214
14	68.5	10.1	294	21	G54042
15	68.5	10.1	350	21	G11213
16	68	10.0	196	20	Y02827
17	68	10.0	252	21	Y73366
18	68	10.0	371	20	Y29193
19	68	10.0	864	20	Y03636
20	67	9.9	455	19	W76439
21	66	9.7	193	21	Y75683
22	66	9.7	343	21	G47999
23	64.5	9.5	185	21	Y28426
24	64	9.4	152	12	R15223
25	64	9.4	287	21	B25784
26	64	9.4	293	21	Y76168
27	64	9.4	305	20	Y29975
28	64	9.4	321	21	B42915
29	63.5	9.4	201	21	G59502
30	63.5	9.4	204	21	G33358
31	63.5	9.4	237	21	G33357
32	63.5	9.4	294	21	G33356
33	63.5	9.4	365	16	R67590
34	63.5	9.4	416	20	Y21591
35	63.5	9.4	531	21	Y75537
36	63	9.3	424	21	Y77281
37	63	9.3	424	21	Y78833
38	62.5	9.2	171	21	G54995
39	62.5	9.2	174	21	G14734
40	62.5	9.2	200	21	G56114
41	62.5	9.2	230	21	G14733
42	62.5	9.2	235	21	G39986
43	62.5	9.2	291	21	G39985
44	62.5	9.2	306	21	G14732
45	62.5	9.2	365	16	R67592

## ALIGNMENTS

RESULT 1	
ID W69225	standard; Protein: 129 AA.
XX AC W69225;	
XX DT 18-FEB-1999	(first entry)
XX DE NADH dehydrogenase subunit NDS-2.	
XX KW NADH dehydrogenase subunit; sympathetic nervous system disease; myopathy; nicotinamide-adenine dinucleotide dehydrogenase; cancer; leukaemia;	
KW immune system disorder; neurodegenerative disease; therapy; NDS-2.	
XX OS Homo sapiens.	
XX PN W09831815-A2.	
XX PD 23-JUL-1998.	
XX PF 17-DEC-1997;	97WO-US23970.
XX PR 17-JAN-1997;	97US-0785065.
XX PA (INCYT-) INCYTE PHARM INC.	
XX PI Bandman O, Goli SK, Hillman JL;	
XX DR WPI: 1998-414112/35.	
XX DR N-PSDB: V44787.	
XX PT Human nicotinamide-adenine dinucleotide dehydrogenase sub:units -	
PT useful for, e.g. diagnosis, treatment and prevention of cancer,	
PT myopathy, immune system disease and neurodegeneration	

XX Claim 19; Fig 2; 80pp; English.

CC This sequence represents the NADH (reduced nicotinamide-adenine  
CC dinucleotide) dehydrogenase subunit, NDS-2, of the invention. Cells  
CC containing the DNA are used to produce the recombinant subunits.  
CC Antagonists of NDS-1 (typically antisense sequences or ribozymes) are  
CC used to treat or prevent cancer (leukaemia and solid cancers) and immune  
CC system disorders (e.g. asthma, diabetes, rheumatoid arthritis,  
CC osteoporosis and many others). NDS-2 and NDS-4 are used to treat myopathy  
CC (e.g. ophthalmoplegia, myoclonic epilepsy and lactic acidosis), while  
CC their antagonists are used to treat cancer and disease of the sympathetic  
CC nervous system (e.g. hypertension, arrhythmia and migraine). NDS-3 is  
CC also used to treat myopathy and its antagonists to treat cancer and  
CC neurodegenerative disease (e.g. Alzheimer's, Huntington's and Parkinson's  
CC diseases, epilepsy and Down's syndrome). In all cases NDS or their  
CC antagonists may be expressed from gene therapy vectors. Ab may be used  
CC therapeutically as antagonist; as immunoassay reagent for diagnosis or  
CC monitoring such diseases; in competitive screening assays for agents that  
CC bind specifically to the subunits, and for affinity purification of the  
CC subunits from natural sources. The DNAs are useful as primers and probes  
CC for diagnosis and monitoring (including detecting predisposition to  
CC cancer); for gene mapping or identifying related sequences, while the  
CC subunits are also used to raise antibodies and to screen for specific  
CC binding agents.

SO Sequence 129 AA;

Query Match 100.0%; Score 678; DB 19; Length 129;  
Best Local Similarity 100.0%; Pred. No. 7.4e-74;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSFPYKSSRLTLEPTDPAEYNISPETRRAORIALIARQLKREYLLQVNDPRRGII 60  
DB 1 msfpykssrltpeclpaeynispeetraaeialraqkreyllqyndpnrrgll 60  
OY 61 ENPALLRNAYARTIWNVFNFRPTPKNSLGMALCGFPLFIYIIKTERDRKELIOEGK 120  
DB 61 enpallrwayartlnvfnfrptpknsimgalcgplfiylikterdrkexldeqk 120  
OY 121 LDRTHLSY 129  
DB 121 ldrthlsy 129

RESULT 2

ID 176629 standard; Protein; 113 AA.

AC Y76629;

DT 10-APR-2000 (first entry)

XX Human ovarian tumor EST fragment encoded protein 125.

KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;

KW gene therapy; treatment.

OS Homo sapiens.

PN DE19817557-A1.

PD 21-OCT-1999.

PF 09-APR-1998; 98DE-1017557.

PR 09-APR-1998; 98DE-1017557.

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

DR WPI; 1999-591920/51.

DR N-PSDB; 277502.

XX New nucleic acid sequences expressed in ovarian, and some other, cancer  
XX tissues, and derived polypeptides, for treatment of ovarian cancer and  
XX identification of therapeutic agents.

PS Claim 25; Page 295; 310pp; German.

CC This invention describes novel nucleic acid (cDNA) sequences (A) which  
CC have anticancer activity and are highly expressed in ovarian tumor  
CC tissue (and some also in testis and breast cancer tissue). The products  
CC of the invention can be used for gene therapy. (A) are used (i) for  
CC recombinant expression of polypeptides (B) and (ii) to isolate complete  
CC genes. (B) are used (i) to identify agents suitable for treatment of  
CC ovarian cancer; (ii) directly for treating this form of cancer (including  
CC expression from gene therapy vectors) and (iii) for generation of  
CC specific antibodies. (A) are identified by assembling ESTs (expressed  
CC sequence tags) from a particular tissue type before comparison of  
CC expression patterns. This allows a significantly longer fragment of the  
CC gene to be revealed, so should reduce the number of failures associated  
CC with the fact that ESTs from different libraries may represent different  
CC parts of the same unknown gene, distorting the estimated frequency of  
CC occurrence in a particular tissue. Y76505-Y76638 represent protein  
CC fragments encoded by the human ovarian tumor CDNA library derived EST  
CC fragments represented in 277450-277572.

SO Sequence 113 AA;

Query Match 74.2%; Score 503; DB 20; Length 113;  
Best Local Similarity 99.0%; Pred. No. 7.2e-53;  
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 34 AERLAIARQLKREYLLQVNDPRRGIIENPALLRNAYARTIWNVFNFRPTPKNSLGMALC 93  
DB 18 serlaraqkreyllqyndpnrrgllienpallrwayartlnvfnfrptpknsimgalc 77  
OY 94 GFGPLFIYIIKTERDRKELIOEGKLDPRHLSY 129  
DB 78 gfgplfiylikterdrkexldeqkldrthlsy 113

RESULT 3

ID B53504 standard; Protein; 16 AA.

AC B53504;

DT 09-MAR-2001 (first entry)

DE Human colon cancer antigen protein sequence SEQ ID NO:1044.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytotoxic; cardioactive; neuroprotective; vulnary;

KW immunomodulatory; muscular; gynaecological; gastrointestinal;

KW nephrotropic; antineoplastic; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder.

OS Homo sapiens.

PN WO200055351-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000WO-US05883.

PR 12-MAR-1999; 99US-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;  
XX WPI; 2000-587534/55.  
DR N-PSDB; C98261.  
XX  
PT Colon cancer associated gene sequences, referred to as colon cancer  
PT antigens, useful for the treatment, prevention, and diagnosis of colon  
PT disorders such as colon cancer  
XX  
PS Claim 11: Page 1624; 2104pp; English.  
XX  
CC C97991 to C98763 encode the human colon cancer associated proteins,  
CC called human colon cancer antigens, given in B53234 to B54006. The human  
CC colon cancer antigens can have cytostatic, cardiocactive, muscular;  
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
CC vulnerrary, nephrotropic, antinefctive and antibacterial activities, and  
CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
CC proteins and antibodies to the proteins are useful for the prevention,  
CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins  
CC may also be used to prevent diseases such as neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, wounds, renal disorders, infectious  
CC diseases, and cardiovascular disorders. C98764 to C98772 and B54007  
CC represent sequences used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 16 AA:  
  
Query Match 11.2%; Score 76; DB 21; Length 16;  
Best Local Similarity 93.8%; Pred. No. 0.013;  
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 114 KLTQEGKLDRTFHLST 129  
Db 1 KLTQEGKLDRTFHLST 16  
  
RESULT 4  
G33365 633365 standard; Protein; 290 AA.  
XX  
AC G33365;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
XX Zea mays protein fragment SEQ ID NO: 40416.  
XX  
KM Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0144624.  
PR 15-JUL-1999; 99US-0144605.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.

PR	14-Oct-1999;	99US-0159331.
PR	14-Oct-1999;	99US-0159637.
PR	14-Oct-1999;	99US-0159638.
PR	18-Oct-1999;	99US-0159584.
PR	21-Oct-1999;	99US-0160741.
PR	21-Oct-1999;	99US-0160767.
PR	21-Oct-1999;	99US-0160768.
PR	21-Oct-1999;	99US-0160770.
PR	21-Oct-1999;	99US-0160814.
PR	21-Oct-1999;	99US-0160815.
PR	22-Oct-1999;	99US-0160980.
PR	22-Oct-1999;	99US-0160981.
PR	22-Oct-1999;	99US-0160989.
PR	25-Oct-1999;	99US-0161404.
PR	25-Oct-1999;	99US-0161405.
PR	25-Oct-1999;	99US-0161406.
PR	26-Oct-1999;	99US-0161359.
PR	26-Oct-1999;	99US-0161360.
PR	26-Oct-1999;	99US-0161361.
PR	28-Oct-1999;	99US-0161920.
PR	28-Oct-1999;	99US-0161992.
PR	28-Oct-1999;	99US-0161993.
PR	29-Oct-1999;	99US-0162142.

```

QY      4 PKKPSLSLTLEETLDPPEVNSIPETRAOAEERLAIAPLAKREYLLQYN -PPNR----- 57
Db      10 pLlCPpLvarp----pa--ssppagrpasqrlararackidamgyrkadpilaags 63
QY      58 -----GLINPALLRWAAYARTINYPNR-----PPKKSLSMGALCGEPL 98
Db      64 sILypmtespe-lrwatvkrIyIltvqlamCaavsatfvkvavsnlfvssnagIaly 122
QY      99 IFT 101
Db      123 111 125

```

23-FEB-2001 (first entry)  
Protein involved in cell cycle regulation SEQ ID 14

*Zea mays.*

WO200065040-A2.

02-NOV-2000.

13-APR-2000; 2000WO-US09975

22-APR-1999; 99US-0130849.

(PION-) PIONEER HI-BRED INT INC.

Helentjaris TG, Habben JE, Sun Y;

WPL; 2000-687333/67.  
N-DEPD; 003104

XX  
DT

nucleic acids useful for producing transgenic plants, preferably maize,

PT With increased cell cycle gene activity, preferably activity of cyclin  
and/or cyclin-dependent kinase.

PS Claim 16; Page 93-94; 122pp; English.

CC Polynucleotide sequences C83101 - C83113 encode proteins B35794 - B35806  
CC which are involved in regulating the cell cycle. The protein and DNA  
CC sequences have been isolated from Zea mays (corn), and the invention also  
CC includes oligonucleotides C83114 - C83139 which are related to the cell  
CC cycle polynucleotides. The cell cycle polynucleotide sequences are  
CC useful for producing transgenic plants such as maize, soybean, sunflower,  
CC sorghum, canola, wheat, alfalfa, cotton, rice, barley and millet with  
CC increased levels of cell cycle gene activity, such as activity of cyclin  
CC and cyclin-dependent kinases. The DNA sequences are also useful as probes  
CC for detecting deficiencies in the level of mRNA in screening for desired  
CC transgenic plants, for detecting mutations in the gene, for monitoring  
CC upregulation of expression or changes in enzyme activity in screening  
CC assays of compounds, for detecting any number of allelic variants,  
CC orthologs or paralogues of the gene, and site-directed mutagenesis in  
CC eukaryotic cells. The DNA sequences are also useful for recombinant  
CC expression of the encoded polypeptides and as immunogens for preparing  
CC and screening antibodies. A transgenic plant comprising an expression  
CC cassette including a cell cycle regulatory gene is useful for assaying  
CC enzyme agonists and antagonists, and as immunogens or antigens to obtain  
CC antibodies. The antibodies are useful in assaying expression levels of  
CC cell cycle regulatory proteins, for identifying and isolating nucleic  
CC acids from expression libraries, for identifying homologues of  
CC polypeptides from other species, and for purification of the proteins.

CC Sequence 292 AA;

Query Match 10.3%; Score 69.5; DB 21; Length 292;

Best Local Similarity 34.0%; Pred. No. 3.9; Mismatches 1; Gaps 1;

Matches 16; Conservative 12; Indels 18; Indels 1; Gaps 1;

OY 2 SPPKYPSSLRLPTLPDPAEYN-ISPETRRRAQERLAIRAOQLKREY 47

DB 239 tfpkwpvdiatvptlpsgidliskmlrldpkskrtaraalehdy 285

RESULT 6

W95690 W95690 standard; Protein; 294 AA.

XX AC W95690;

XX DT 08-JUN-1999 (first entry)

XX DE Oryza sativa Cdc2 protein kinase.

XX KW cdc2; protein kinase; diagnosis; infection; immunosuppression;

XX KM AIDS; acquired immunodeficiency syndrome; diagnosis; therapy;

XX OS Oryza sativa.

XX PN W09856799-A1.

XX PD 17-DEC-1998.

XX PF 12-JUN-1998; 98WO-US12100.

XX PR 13-JUN-1997; 97US-0874347.

XX PA (MAYO-) MAYO FOUNDATION.

XX PI Gustafson MP, Leof EB, Limper AH, Thomas CF;

XX DR WPI; 1999-080877/07.

XX PT New isolated polynucleotide encoding a Pneumocystis carinii Cdc2  
PT polypeptide - useful for diagnosing or monitoring P. carinii

PT Infection in patients with chronic immunosuppression  
XX Example; Pages 44-45; 67pp; English.

CC The sequence is that of a Cdc2 polypeptide.  
CC The Cdc2 polypeptides are useful in treating P. carinii  
CC pneumonia in patients with chronic immunosuppression. Conditions  
CC associated with P. carinii include acquired immune deficiency syndrome  
CC (AIDS), solid tumors, hematological malignancies, organ transplantation,  
CC and inflammatory conditions. The polypeptides are additionally useful  
CC in identifying agents that inhibit the phosphorylation activity of  
CC P. carinii Cdc2 polypeptides. The polypeptide and its antibodies are  
CC useful for diagnosing or monitoring P. carinii infection in patients.  
CC The P. carinii Cdc2 polypeptides allow a therapeutic approach to treating  
CC P. carinii pneumonia because they are not limited by significant side  
CC effects. Polynucleotides encoding P. carinii Cdc2 enable in vitro  
CC production of the protein, which can be used in studying the organism  
CC (especially its life cycle) and developing new therapies.

CC Sequence 294 AA;

Query Match 10.3%; Score 69.5; DB 20; Length 294;

Best Local Similarity 36.2%; Pred. No. 3.9; Mismatches 20; Indels 1; Gaps 1;

Matches 17; Conservative 9; Indels 20; Indels 1; Gaps 1;

OY 2 SPPKYPSSLRLPTLPDPAEYN-ISPETRRRAQERLAIRAOQLKREY 47

DB 240 atpkwqadlptvptlpsgidliskmlrldpkskrtaraalehdy 286

RESULT 7

R88358 R88358 standard; Protein; 656 AA.

XX AC R88358;

XX DT 15-OCT-1995 (first entry)

XX DE Human methylene-tetrahydrofolate-reductase.

XX KM Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;

XX KW cardiovascular disease; neurological disease; folic acid metabolism;

XX OS Homo sapiens.

XX PN W09533054-A1.

XX PD 07-DEC-1995.

XX PF 25-MAY-1995; 95WO-CA00314.

XX PR 26-MAY-1994; 94GB-0010620.

XX PA (UYMC-) UNIV MCGILL.

XX PI Goyette P, Rozen R;

XX DR WPI; 1996-030565/03.

XX DR N-PSDB; T09689.

PT Human methylene-tetrahydrofolate reductase cDNA probe - for  
PT detection of sequence abnormalities in methylene-tetrahydrofolate  
PT reductase e.g. in cardiovascular, neurological or folic acid  
PT metabolism disorders

PS Disclosure; Fig. 6A-6C; 66pp; English.

CC This is the protein sequence encoding human MTHFR, the gene of  
CC which has been localised to chromosome 1p36.3. Deficiencies of  
CC this protein may lead to cardiovascular and neurological disorders  
CC and disorders influences by folic acid metabolism.

XX Sequence 656 AA:

Query Match 10.2%; Score 69; DB 17; Length 656;  
Best Local Similarity 28.7%; Pred. No. 13;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

QY 35 ERLAIRAQLKREYLLQYDNPGRGLIENPALLRMAYARTINYPNRPPTKNSLMGALCG 94  
DB 458 eplaetaelikeellrv---nrgill-----tinsgpnngkpsd---plvg 499  
QY 95 FGP---LIF---IYIITKTERDRKEKLIQ---EGKLDRTFHL 127  
DB 500 wgpssgyvfqkayleftrstaallqvkkylrnyhl 540

RESULT 8  
R88359  
ID R88359 standard; Protein; 660 AA.

AC R88359;  
DT 15-OCT-1995 (first entry)

DE Human methylene-tetrahydrofolate-reductase cDNA.

KW Methylene-tetrahydrofolate-reductase; MTHFR; gene therapy;  
KW cardiovascular disease; neurological disease; folic acid metabolism;  
KW EC-1.5.1.20; enzyme.

OS Homo sapiens.

PN WO9533054-A1.

PD 07-DEC-1995.

PF 25-MAY-1995; 95MO-CA00314.

PR 26-MAY-1994; 94GB-0010620.

PA (UYMC-) UNIV MCGILL.

PI Goyette P, Rozen R;

DR WPI: 1996-030565/03.

DR N-PSDB: T09694.

PT Human methylene-tetrahydrofolate reductase cDNA probe - for  
PT detection of sequence abnormalities in methylene-tetrahydrofolate  
PT reductase e.g. in cardiovascular, neurological or folic acid  
PT metabolism disorders

PS Disclosure: Fig. 1A-1F; 66pp; English.

CC The gene encoding this human MTHFR protein has been localised to  
CC chromosome 1p36.3. Deficiencies of this protein may lead to  
CC cardiovascular and neurological disorders and disorders influences  
CC by folic acid metabolism.

XX Sequence 660 AA:

Query Match 10.2%; Score 69; DB 17; Length 660;  
Best Local Similarity 28.7%; Pred. No. 13;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

QY 35 ERLAIRAQLKREYLLQYDNPGRGLIENPALLRMAYARTINYPNRPPTKNSLMGALCG 94  
DB 462 eplaetaelikeellrv---nrgill-----tinsgpnngkpsd---plvg 503  
QY 95 FGP---LIF---IYIITKTERDRKEKLIQ---EGKLDRTFHL 127

DB 504 wgpssgyvfqkayleftrstaallqvkkylrnyhl 544

RESULT 9

ID Y96186  
Y96186 standard; Protein; 660 AA.

AC Y96186;

DT 19-DEC-2000 (first entry)

DE Human methylenetetrahydrofolate reductase.

KW Methylene-tetrahydrofolate reductase; MTHFR; human; folic acid;  
KW cardiovascular disorder; cancer; neuroblastoma;  
KW colorectal carcinoma; osteoporosis; neural tube defect;  
KW neurological disorder; gene therapy; diagnosis.

OS Homo sapiens.

PN WO200052205-A2.

PD 08-SEP-2000.

PF 28-FEB-2000; 2000MO-IB00442.

PR 01-MAR-1999; 99US-0258928.

PA (UYMC-) UNIV MCGILL.

PI Rozen R, Goyette P;

DR WPI: 2000-572192/53.

DR N-PSDB: A50633, A50634.

PT cDNA probe for the human methylenetetrahydrofolate reductase (MTHFR),  
PT useful in gene therapy and for diagnosing or treating MTHFR deficiency  
PT which is associated with cardiovascular disorders or cancer  
PS Claim 1; Fig 1A-F; 93pp; English.

CC The present sequence is that of human methylenetetrahydrofolate  
CC reductase (MTHFR), an enzyme catalysing the NADPH-linked reduction  
CC of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, a  
CC co-substrate for methylation of homocysteine to methionine. A cDNA  
CC probe for human MTHFR is claimed. This probe can be used to  
CC identify MTHFR sequence abnormalities in individuals with severe or  
CC mild MTHFR deficiency. MTHFR deficiencies may be associated with  
CC a cardiovascular disorder, cancer (especially neuroblastoma or  
CC colorectal carcinoma), osteoporosis, neural tube defect in an  
CC offspring of a patient, neurological disorders, and other disorders  
CC influenced by folic acid metabolism. Also claimed are methods for  
CC treating MTHFR deficiency by gene therapy or by administration of  
CC MTHFR protein. Cancer can be treated by inhibiting MTHFR gene  
CC expression or MTHFR protein activity, or by administering an agent  
CC that modifies MTHFR gene expression.

XX Sequence 660 AA:

Query Match 10.2%; Score 69; DB 21; Length 660;  
Best Local Similarity 28.7%; Pred. No. 13;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

QY 35 ERLAIRAQLKREYLLQYDNPGRGLIENPALLRMAYARTINYPNRPPTKNSLMGALCG 94  
DB 462 eplaetaelikeellrv---nrgill-----tinsgpnngkpsd---plvg 503  
QY 95 FGP---LIF---IYIITKTERDRKEKLIQ---EGKLDRTFHL 127  
DB 504 wgpssgyvfqkayleftrstaallqvkkylrnyhl 544



PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 27-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151338.  
 PR 07-SEP-1999; 99US-0151330.  
 PR 10-SEP-1999; 99US-0152363.  
 PR 13-SEP-1999; 99US-0153070.  
 PR 15-SEP-1999; 99US-0153758.  
 PR 16-SEP-1999; 99US-0154018.  
 PR 20-SEP-1999; 99US-0154039.  
 PR 22-SEP-1999; 99US-0154779.  
 PR 23-SEP-1999; 99US-0155139.  
 PR 24-SEP-1999; 99US-0155486.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 29-SEP-1999; 99US-0156458.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158233.  
 PR 12-OCT-1999; 99US-0158233.  
 PR 13-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 14-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 22-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 25-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 204;  
 Best Local Similarity 36.2%; Pred. No. 3.2; Indels 1; Gaps 1;  
 Matches 17; Conservative 8; Mismatches 21;

OY 2 SPFKPKSSLRTPLETLPAETN-ISPETRAQAERLAIRAOIKREY 47  
 DB 150 atpkkwptdletfvpnlbdpdydliskmlmdpkrtrlnaraalehy 196

RESULT 11  
 ID G11215  
 XX G11215 standard; Protein; 237 AA.  
 AC G11215;  
 XX 17-QCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 9849.  
 DE Protein identification: signal transduction pathway; metabolic pathway;  
 KW hybridisation assay: genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX Arabidopsis thaliana.  
 OS EPI033405-A2.  
 FN  
 XX 06-SEP-2000.  
 PD  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 24-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 27-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136392.  
 PR 01-JUN-1999; 99US-0136782.  
 PR 03-JUN-1999; 99US-0137222.  
 PR 04-JUN-1999; 99US-0137328.  
 PR 07-JUN-1999; 99US-0137502.  
 PR 08-JUN-1999; 99US-0137724.  
 PR 10-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139219.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 18-JUN-1999; 99US-0139763.



PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140354.  
PR 28-JUN-1999; 9905-0140623.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142603.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144684.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 18-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150565.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.

PR 27-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155133.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0158293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160768.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161922.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 237;  
Best Local Similarity 36.2%; Pred. No. 3.9;  
Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

QY 2 SPKKYPSLRTLPETDAEYN-ISPETRAQAEIRAIQAOLKREY 47  
Db 183 atpkxpkdletfvpnlidpdydliskmlmpktrinaaraalehey 229

RESULT 12  
G54043  
ID G54043 standard; Protein; 237 AA.

XX G54043;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 68864.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.  
XX EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 08-APR-1999; 99US-0128234.  
PR 16-APR-1999; 99US-0128714.  
PR 19-APR-1999; 99US-0129845.  
PR 11-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 28-APR-1999; 99US-0130510.  
PR 30-APR-1999; 99US-0130891.  
PR 05-MAY-1999; 99US-0131449.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0138458.  
PR 18-JUN-1999; 99US-0138459.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139751.  
PR 21-JUN-1999; 99US-0139753.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 29-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 13-JUL-1999; 99US-0142977.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145921.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 10-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 07-SEP-1999; 99US-0151930.  
PR 10-SEP-1999; 99US-0152363.  
PR 13-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 237;  
Best Local Similarity 36.2%; Pred. No. 3.9;  
Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

OY 2 SEPKYPSLRITPEIDPAEYN-ISPETRRQAERLAIRAQIKREY 47  
DI 183 afpkwptdletfvpnlpdpgvdllskmlmdpklrinaralehey 229

RESULT 13

ID G11214 standard; Protein; 294 AA.

AC G11214;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9848.

KW Protein identification: signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0133548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132485.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134766.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 17-JUL-1999; 99US-0144086.  
 PR 18-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144332.  
 PR 20-JUL-1999; 99US-0144332.  
 PR 20-JUL-1999; 99US-0144332.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 23-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 26-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 09-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 10-AUG-1999; 99US-0147935.  
 PR 11-AUG-1999; 99US-0148171.  
 PR 12-AUG-1999; 99US-0148319.  
 PR 13-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 15-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.  
 PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 23-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 25-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151330.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154039.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 22-SEP-1999; 99US-0154779.  
 PR 23-SEP-1999; 99US-0155139.  
 PR 24-SEP-1999; 99US-0155486.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 99US-0156458.

PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158232.  
 PR 13-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159337.  
 PR 14-OCT-1999; 99US-0159337.  
 PR 18-OCT-1999; 99US-0159638.  
 PR 21-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 29-OCT-1999; 99US-0161993.  
 PR 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 294;  
 Best Local Similarity 36.2%; Pred. No. 5.1;  
 Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

OY 2 SFEKYNSSRLTLPETIDPAEYN-ISPETRRQAERLAIKQAKKEY 47  
 DB 240 afpkwkpdlctfvpnlpdgydliskmlndpkrlnaraalehey 286

RESULT 14  
 G54042  
 ID G54042 standard; Protein; 294 AA.  
 AC G54042;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 68863.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 OS Arabidopsis thaliana.  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142290.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144881.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145199.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145931.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148688.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0150566.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155138.  
PR 23-SEP-1999; 99US-0155138.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157733.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 23-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10 1%; Score 68.5; DB 21; Length 294;  
Best Local Similarity 36.2%; Pred. No. 5.1;

Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;

Qy 2 SEPKKXSSSLRTLPETLDPAEVN-ISPETRAQAERLAIRAQIKREY 47  
Db 240 atpkwptdletfvpdpdydlkskmlmpktrlnaraalehey 286

RESULT 15

ID G11213 standard; Protein: 350 AA.

AC G11213;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 9847.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW Hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144087.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 23-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 10.1%; Score 68.5; DB 21; Length 350;  
Best Local Similarity 36.2%; Pred. NO. 6.5;  
Matches 17; Conservative 8; Mismatches 21; Indels 1; Gaps 1;  
QY 2 SFPKPPSLRTPETLDPAEYN-ISPTERRAQARLAIHQKREY 47  
Db 296 afpwxpdlletfvpndpbdgvdlskmlmdpkrlnaraalehey 342

Search completed: April 29, 2001, 06:45:09  
Job time: 625 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: April 29, 2001, 06:40:49 ; Search time 12.61 Seconds  
(Without alignments)  
196.527 Million cell updates/sec

Title: US-09-726-899-3  
Perfect score: 678  
Sequence: 1 MSFPRKPSRLTLPETLDP.....DRKELIOGKIDRPHLSY 129

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

T number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/BackfileSI.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	678	100.0	129	2	US-08-785-065-3
2	514	75.8	129	2	US-08-785-065-10
3	69.5	10.3	294	2	US-08-874-347-26
4	69.5	10.3	294	3	US-09-093-522-26
5	69	10.2	656	3	US-08-738-000-4
6	69	10.2	660	3	US-08-738-000-2
7	64	9.4	152	1	US-07-644-372-2
8	64	9.4	305	2	US-08-946-528-1
9	63	9.3	3898	4	US-08-750-717-2
10	62.5	9.2	366	2	US-08-605-106-11
11	62	9.1	214	4	US-09-587-066-6
12	61.5	9.1	60	1	US-08-370-325-32
13	61.5	9.1	60	1	US-08-461-859-32
14	61.5	9.1	60	5	PCT-US93-10069-32
15	61.5	9.1	353	2	US-08-176-620A-14
16	61.5	9.1	353	2	US-08-461-985-14
17	61.5	9.1	484	2	US-08-836-620A-13
18	61.5	9.1	485	2	US-08-836-620A-2
19	61	9.0	420	2	US-08-846-762-73
20	61	9.0	619	1	US-08-472-934-4
21	61	9.0	619	1	US-08-472-934-12
22	61	9.0	619	2	US-08-323-460A-4
23	61	9.0	619	2	US-08-461-146C-4
24	61	9.0	619	2	US-08-461-146C-12
25	61	9.0	619	3	US-08-461-145C-4
26	61	9.0	619	3	US-08-461-145C-12
27	61	9.0	915	2	US-08-480-917-2

28	61	9.0	1213	2	US-08-937-102-2	Sequence 2, Appl
29	61	9.0	1261	1	US-08-764-100-26	Sequence 26, Appl
30	60.5	8.9	379	3	US-08-622-277A-8	Sequence 8, Appl
31	59.5	8.8	362	1	US-08-464-523B-32	Sequence 32, Appl
32	59.5	8.8	362	2	US-08-440-845D-6	Sequence 6, Appl
33	59.5	8.8	362	4	US-08-868-458-6	Sequence 6, Appl
34	59.5	8.8	716	2	US-08-766-982-1	Sequence 1, Appl
35	59.5	8.8	793	1	US-08-015-985-3	Sequence 3, Appl
36	59.5	8.8	1382	3	US-09-057-570-4	Sequence 4, Appl
37	59	8.7	262	2	US-08-719-697-10	Sequence 10, Appl
38	59	8.7	325	1	US-08-142-473A-9	Sequence 9, Appl
39	59	8.7	325	1	US-08-469-203A-9	Sequence 9, Appl
40	59	8.7	325	1	US-08-469-203A-9	Sequence 9, Appl
41	59	8.7	385	2	US-08-605-106-14	Sequence 14, Appl
42	59	8.7	565	2	US-08-892-770-6	Sequence 6, Appl
43	59	8.7	571	2	US-08-892-770-5	Sequence 5, Appl
44	59	8.7	673	4	US-09-078-347A-2	Sequence 2, Appl
45	58.5	8.6	3033	1	US-07-925-695-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-785-065-3  
: Sequence 3, Application US/08785065  
: Patent No. 5814451  
: GENERAL INFORMATION:  
: APPLICANT: Bandman, Olga  
: APPLICANT: Goli, Surya K.  
: APPLICANT: Hillman, Jennifer L.  
: TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE  
: NUMBER OF SEQUENCES: 12  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Incyte Pharmaceuticals, Inc.  
: STREET: 3174 Porter Drive  
: CITY: Palo Alto  
: STATE: CA  
: COUNTRY: USA  
: ZIP: 94304  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Diskette  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: DOS  
: SOFTWARE: FASTSEQ for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/785,065  
: FILING DATE: Herewith  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER:  
: FILING DATE:  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Billings, Lucy J.  
: REGISTRATION NUMBER: 36,749  
: REFERENCE/DOCKET NUMBER: PF-0187 US  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415-855-0555  
: TELEFAX: 415-845-4166  
: TELEX:  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 129 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: IMMEDIATE SOURCE:  
: LIBRARY: Consensus  
: CLONE: Consensus  
: US-08-785-065-3  
Query Match 100.0% ; Score 678 ; DB 2 ; Length 129 ;

Best Local Similarity 100.0%; Pred. No. 8.7e-75;  
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSFRRKPPSSLRTPETLPDAEYNIISPETRRQAERLAIRAQLKREYLLQYNDPNRRGLI 60

Db 1 MSFRRKPPSSLRTPETLPDAEYNIISPETRRQAERLAIRAQLKREYLLQYNDPNRRGLI 60

QY 61 ENPALLRMAYARTINVPNFRPTPKNSLMGALCGFGPLIFYYIIKTERDRKREKLIQEGK 120

Db 61 ENPALLRMAYARTINVPNFRPTPKNSLMGALCGFGPLIFYYIIKTERDRKREKLIQEGK 120

QY 121 LDRTEHLISY 129

Db 121 LDRTEHLISY 129

RESULT 2  
US-08-785-065-10

Sequence 10, Application US/08785065

Patent No. 5814451

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Goll, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL SUBUNITS OF NADH DEHYDROGENASE

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/785,065

FILING DATE: Herewith

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0187 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 129 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: Genbank

CLONE: 114

US-08-785-065-10

Query Match 75.8%; Score 514; DB 2; Length 129;

Best Local Similarity 73.6%; Pred. No. 6.5e-55;

Matches 95; Conservative 18; Mismatches 16; Indels 0; Gaps 0;

QY 1 MSFRRKPPSSLRTPETLPDAEYNIISPETRRQAERLAIRAQLKREYLLQYNDPNRRGLI 60

Db 1 MSFRRKPPSSLRTPETLPDAEYNIISPETRRQAERLAIRAQLKREYLLQYNDPNRRGLI 60

QY 61 ENPALLRMAYARTINVPNFRPTPKNSLMGALCGFGPLIFYYIIKTERDRKREKLIQEGK 120

Db 61 ENPALLRMAYARTINVPNFRPTPKNSLMGALCGFGPLIFYYIIKTERDRKREKLIQEGK 120

QY 121 LDRTEHLISY 129

Db 121 LDRTEHLISY 129

RESULT 3  
US-08-874-347-26

Sequence 26, Application US/08874347

Patent No. 5863741

GENERAL INFORMATION:

APPLICANT: Limper, Andrew H.

APPLICANT: Leof, Edward B.

APPLICANT: Thomas, Charles F.

APPLICANT: Gustafson, Michael P.

TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C., P.A.

STREET: 60 South Sixth Street, Suite 3300

CITY: Minneapolis

STATE: MN

COUNTRY: USA

ZIP: 55402

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/874,347

FILING DATE: 13-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Ellinger, Mark S.

REGISTRATION NUMBER: 34,812

REFERENCE/DOCKET NUMBER: 07039/055001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612-335-5070

TELEFAX: 612-288-9696

TELEX:

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 294 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-874-347-26

Query Match 10.3%; Score 69.5; DB 2; Length 294;

Best Local Similarity 36.2%; Pred. No. 1.4;

Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

QY 2 SEPKKPPSSLRTPETLPDAEYNIISPETRRQAERLAIRAQLKREY 47

Db 240 APPKQADLATIVPTLPDPAAGLDLSKMLRYPNKRITRQALHEHY 286

RESULT 4  
US-09-093-522-26

Sequence 26, Application US/09093522

Patent No. 6015700

GENERAL INFORMATION:

APPLICANT: Limper, Andrew H.

APPLICANT: Leof, Edward B.

APPLICANT: Thomas, Charles F.

APPLICANT: Gustafson, Michael P.

TITLE OF INVENTION: CDC2 PROTEIN KINASE FROM PNEUMOCYSTIS  
TITLE OF INVENTION: CARINITI  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C., P.A.  
STREET: 60 South Sixth Street, Suite 3300  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/093,522  
FILING DATE: 08-JUN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/874,347  
FILING DATE: 13-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ellinger, Mark S.  
REGISTRATION NUMBER: 34,812  
REFERENCE/DOCKET NUMBER: 07039/055002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-335-5070  
TELEFAX: 612-288-9696  
TELEX:  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 294 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-093-522-26

Query Match 10.3%; Score 69.5; DB 3; Length 294;  
Best Local Similarity 36.2%; Pred. No. 1.4;  
Matches 17; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 2 SEPKKPSRLTLPETLDAEYN-ISPEFRQAQERLAIRPAOLKREY 47  
Db 240 APPKQADLAVITPLDPAGLDLKMLKPEPKRITANQALEHEY 286

US-08-738-000-4  
Sequence 4, Application US/08738000  
Patent No. 6074821  
GENERAL INFORMATION:  
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
REDUCTASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: Continental Plaza - 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,000  
FILING DATE:  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/CA95/00314  
FILING DATE: 25-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9410620.0  
FILING DATE: 26-MAY-1994  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 656 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-738-000-4

Query Match 10.2%; Score 69; DB 3; Length 656;  
Best Local Similarity 28.7%; Pred. No. 4.8;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

OY 35 ERLAIRAQLKREYLLQYNDPNRGLIENPALLRMAYARTINYPNRPKSLMGALCG 94  
Db 458 EPLAETSLKELLRV---NRGIL-----TINSOPNNGKRPSSD---PIVG 499

OY 95 FGP---LIF--IYYIKTERDRKELIQ---EGKLDRTFHL 127  
Db 500 WPGSGYVFOKAYLEFITSRETAELALLQYLKREYELRVNYHL 540

RESULT 6  
US-08-738-000-2  
Sequence 2, Application US/08738000  
Patent No. 6074821  
GENERAL INFORMATION:  
TITLE OF INVENTION: CDNA FOR HUMAN METHYLENETETRAHYDROFOLATE  
REDUCTASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KLAUBER & JACKSON  
STREET: Continental Plaza - 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/738,000  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/CA95/00314  
FILING DATE: 25-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9410620.0  
FILING DATE: 26-MAY-1994  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 660 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-738-000-2

Query Match 10.2%; Score 69; DB 3; Length 660;  
Best Local Similarity 28.7%; Pred. No. 4.8;  
Matches 29; Conservative 14; Mismatches 32; Indels 26; Gaps 6;

OY 35 ERLAIRAQLKREYLLQYNDPNRGLIENPALLRMAYARTINYPNRPKSLMGALCG 94  
Db 458 EPLAETSLKELLRV---NRGIL-----TINSOPNNGKRPSSD---PIVG 499

Db 462 EPLAETSLKEELRV---NKGIL-----TINSOPNKGKPSD---PIVG 503  
QY 95 FGP---LIF--IYIKTRDRKREKLIQ---EGKIDRTFHL 127  
Db 504 MGPSSGYVOKAVLEFTRSTETAEALLQVLKKYELRVNHL 544

## RESULT 7

US-07-644-372-2  
Sequence 2, Application US/07644372  
Patent No. 5416009  
GENERAL INFORMATION:

APPLICANT: Lazzeri, Mario E.  
APPLICANT: Nulman, Thomas B.  
APPLICANT: Weiss, Nikolaus  
TITLE OF INVENTION: A DNA SEGMENT ENCODING A SPECIFIC  
TITLE OF INVENTION: IMMUNODIAGNOSTIC ANTIGEN  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:

ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1615 L. Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/644,372  
FILING DATE: 19910123  
CLASSIFICATION: 435  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)861-3000  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-07-644-372-2

## Query Match

Local Similarity 9.4%; Score 64; DB 1; Length 152;  
hes 26; Conservative 16; Mismatches 53; Indels 22; Gaps 3;

QY 8 PSSLRTPETLDPAEYN-----ISPEYRAQAERLAIRAOLKREYLLQYNDPNRRG 58  
Db 44 PDVYSTAPTKLVVSYNNLTVMNGNELFTQVKNQPTKVSMDAEPGALYTLVMTDPDAPS 103  
QY 59 LIENPALLMWAVARTINYPNFRPTPKNSLMGALCGFGLIFITYITITERDRKKKL 115  
Db 104 R-KNPVFRFHMHLINI-----SGQNVSSGTVLSDYIRSTKRRHRTSSL 147

## RESULT 8

US-08-946-528-1  
Sequence 1, Application US/08946528  
Patent No. 5958746  
GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
TITLE OF INVENTION: Corley, Neil C.  
TITLE OF INVENTION: ELECTRON TRANSPORT PROTEINS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.

CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,528  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0406 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 305 amino acids  
TYPE: amino acid  
STRANDEDNESS: Single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSNOT16  
CLONE: 1709102  
US-08-946-528-1

## Query Match

Best Local Similarity 9.4%; Score 64; DB 2; Length 305;  
Matches 34; Conservative 15; Mismatches 39; Indels 52; Gaps 7;

QY 8 PSSLRTPETLDPAEYNISPEYRAQAERLA-----IRAOLK----- 44  
Db 149 PSGLLTY---TGKGFHFIQPNKKSPPRRVAKKLGIMAGTGITPMQLIRAILKVEDEP 205  
QY 45 -----REYL--IQYNDPNRRGL---IENPALLMWAVAR---TINYYPN 79  
Db 206 TQCFLEFNQTEKDIILREDELEQLARYPNRFKLMFTLDHP-KDMAYSKGFYADMIRE 264  
QY 80 FRPTPKNSLMGALCGFGLI 99  
Db 265 HLPAPGDDVILLCGPPPMV 284

## RESULT 9

US-08-750-717-2  
Sequence 2, Application US/08750717  
Patent No. 6180109  
GENERAL INFORMATION:

APPLICANT: MOORMANN, Robertus J. M.  
APPLICANT: VAN RIJN, Petrus A.  
TITLE OF INVENTION: Nucleotide Sequences of Pestivirus  
TITLE OF INVENTION: Strains, Polypeptides Encoded by These Sequences and Use  
TITLE OF INVENTION: Thereof for Diagnosis and Prevention of Pestivirus  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

ADDRESSEE: YOUNG & THOMPSON  
STREET: 745 South 23rd Street  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/750,717  
FILING DATE: 24-DEC-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94201743.5  
FILING DATE: 17-JUN-1994  
Prior APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/NL95/00214  
FILING DATE: 16-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: PATCH, Andrew J.  
REGISTRATION NUMBER: 32,925  
REFERENCE/DOCKET NUMBER: BO 39123  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-521-2297  
TELEFAX: 703-685-0573  
TELEX: 248425 EMBON  
FORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3898 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-750-717-2

Query Match 9.3%; Score 63; DB 4; Length 3898;  
Best Local Similarity 23.0%; Pred. No. 2.8e+02;  
Matches 32; Conservative 19; Mismatches 44; Indels 44; Gaps 6;

QY 2 SFPRKSSSLRTLPETDPAEYINISPEPTRAQAEIRLAIRAOQKREYL-LQINDPFRKGLI 60  
DB 2214 SYNYFPLNARKLDDVPPVY-----ATEDEDLAV-----ELGLDMPDPNCGTV 2260  
QY 61 ENDALEMAVARTINVPNRPFPKNSLMGALCGFG-----PLIFIIYIIKTERD 110  
DB 2261 ETGRALQVVGLS-----TAENMLVALFGYGVQALSKRHIPVVDIYSIEDHR- 2310  
QY 111 RKEKLIOEGLDRTFHLIS 129  
DB 2311 -----LEDTHLOY 2319

RESULT 10  
US-08-750-717-2  
Sequence 11, Application US/08605106  
Patent No. 5910631  
GENERAL INFORMATION:  
APPLICANT: Topfer, R.  
APPLICANT: Martini, N.  
TITLE OF INVENTION: MEDIUM CHAIN-SPECIFIC THIOESTERS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.  
STREET: P.O. Box 2938  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/605,106  
FILING DATE: 23-SEPT-1996  
CLASSIFICATION: 800

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP94/02935  
FILING DATE: 01-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Woessner, Warren D.  
REGISTRATION NUMBER: 30,440  
REFERENCE/DOCKET NUMBER: 235,001U51  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-373-6900  
TELEFAX: 612-339-3061  
TELEX:  
FORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 366 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-605-106-11

Query Match 9.2%; Score 62.5; DB 2; Length 366;  
Best Local Similarity 36.1%; Pred. No. 13;  
Matches 13; Conservative 11; Mismatches 11; Indels 1; Gaps 1;

QY 1 MSFPRKSSSLRTLPETDPAEYINISPEPTRAQAE 35  
DB 225 LAFPEENNSSIKRKPLEDDPAQYSMLEKRRRDD 260

RESULT 11  
US-09-587-066-6  
Sequence 6, Application US/09587066  
Patent No. 6210945  
GENERAL INFORMATION:  
APPLICANT: LUNNEN, KEITH D.  
APPLICANT: MORGAN, RICHARD D.  
APPLICANT: MEISEL, TIMOTHY  
APPLICANT: WILSON, GEOFFREY G.  
TITLE OF INVENTION: METHOD FOR CLONING AND PRODUCING THE RsaI RESTRICTION  
TITLE OF INVENTION: ENDONUCLEASE IN E. COLI AND PURIFICATION OF THE  
FILE REFERENCE: NEB-179  
CURRENT APPLICATION NUMBER: US/09/587,066  
CURRENT FILING DATE: 2000-06-02  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6  
LENGTH: 214  
TYPE: PRP  
ORGANISM: rhodospseudomonas sphaeroides  
US-09-587-066-6

Query Match 9.1%; Score 62; DB 4; Length 214;  
Best Local Similarity 29.1%; Pred. No. 7.5;  
Matches 23; Conservative 10; Mismatches 38; Indels 8; Gaps 3;

QY 11 LRTLPETDPAEYINISPEPTRAQAEIRLAIRA-----QIKREYLLQYNDPFRKGLIENPA 64  
DB 83 IATIPHL-PANNSIVAARAGYPRDLFVSGATRHCEFKATSNMGGDDBRRVLTSAPT 141  
QY 65 -LLRAVARTINVPNRP 82  
DB 142 KMIRLVNSRQGVAPNHP 160

RESULT 12  
US-08-370-225-32  
Sequence 32, Application US/08370225  
Patent No. 5580736  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger

APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/370,225  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-370-225-32

Query Match 9.1%; Score 61.5; DB 1; Length 60;  
Best Local Similarity 29.1%; Pred. No. 1.6;  
Matches 16; Conservative 7; Mismatches 17; Indels 15; Gaps 2;

QY 3 FPKYPPSSIRLTPETL--DPAEYNIISPETRRQAQERLAIRAQKREYLYQYNDPN 55  
DB 13 FPRVNPKGIDLQRMVLVDP-----AKRITKALEHPYLYQYHDPN 54

RESULT 13  
US-08-461-859-32  
Sequence 32, Application US/08461859  
Patent No. 5786169  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,859  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/370,225  
FILING DATE: January 9, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: October 30, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lech, Karen F.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/143002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-461-859-32

Query Match 9.1%; Score 61.5; DB 1; Length 60;  
Best Local Similarity 29.1%; Pred. No. 1.6;  
Matches 16; Conservative 7; Mismatches 17; Indels 15; Gaps 2;

QY 3 FPKYPPSSIRLTPETL--DPAEYNIISPETRRQAQERLAIRAQKREYLYQYNDPN 55  
DB 13 FPRVNPKGIDLQRMVLVDP-----AKRITKALEHPYLYQYHDPN 54

RESULT 14  
PCT-US93-10069-32  
Sequence 32, Application PC/TUS9310069  
GENERAL INFORMATION:  
APPLICANT: Brent, Roger  
APPLICANT: Gyuris, Jeno  
APPLICANT: Golemis, Erica  
TITLE OF INVENTION: Interaction Trap System for Isolating  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/10069  
FILING DATE: 20-OCT-1993  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/969,038  
FILING DATE: 10/30/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/143001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154



